WORKFORCE SCHEDULING FOR SHORT LEAD TIME ENVIRONMENT UNDER UNCERTAIN DEMAND

Edith Walschot

September 2022

UNIVERSITY OF TWENTE.

WORKFORCE SCHEDULING FOR SHORT LEAD TIME ENVIRONMENT UNDER UNCERTAIN DEMAND

A Thesis submitted in fulfilment of the requirements for the master's degree of

MSc Industrial Engineering & Management

Production & Logistic Management specialisation
Faculty of Behavioural, Management and Social Sciences
University of Twente
September 2022

Public version

Author
Edith Walschot

Ceva Logistics Heerlen
Earl Bakkenstraat 10
6422 PJ, Heerlen
The Netherlands

University of Twente
Drienerlolaan 5
7522 NB, Enschede
The Netherlands

Company Supervisors
J. Hoop
Regional Project Engineer
BSc. J. van den Berg
Area Engineering Lead & IT Lead

University of Twente Supervisors
Dr. ir. E. Topan
Lead supervisor
Dr. ir. I. Seyran-Topan
Supervisor

UNIVERSITY OF TWENTE.

Preface

This thesis is the last step to finish my Master's Degree in Industrial Engineering and Management at the University of Twente and therefore also marks the end of my student life. During my HBO education it became clear to me that I needed to sharpen my analytical skills. That is why I chose to follow the premaster here at the University. I immediately felt at home and the subsequent Master's matched my interests perfectly. I am very happy with the knowledge and experiences I gained during the Master's and I am looking forward to what the future will bring.

I am very grateful to everyone at Ceva Logistics who was involved in this project. Despite the many operational tasks under high time pressure, everybody managed to make time to provide me with the information needed for this thesis. In addition to the many people who helped with the project, I would like to express my gratitude to Jibbe in particular. His knowledge of the operational processes and the data structure within the WMS system, ensured that I was able to quickly collect the necessary knowledge and data for this research. Also I am grateful for the insights that Jibbe, Jorik and Manfred provided me with. This helped me to match this research to the strategic goals of the site in Heerlen.

From the University of Twente, I would like to thank my first Supervisor Engin Topan. During his classes I got to know Engin as a very committed and enthusiastic teacher. With the same commitment, Engin supervised me during this thesis and from time to time he challenged me to use additional methods to compare the results. His valuable feedback and knowledge helped me raise the quality of the thesis to a higher level. And, I would like to thank Ipek Seyran Topan, who joined towards the end of this Master Thesis, for providing valuable input as my second supervisor.

Lastly, I would like to thank my parents, friends and boyfriend who supported me during this project. My parents taught me to pursue my goals in life and supported me unconditionally to achieve them. I would like to thank my friends during my stay in Twente. Even though we saw each other less often, it was no less fun. I also want to thank Sarah separately, for proofreading this thesis. Lastly, a big thank you to my boyfriend for his patience and support even during long studying nights.

I hope you enjoy reading my master thesis.

Edith Walschot
Heerlen, September 2022

Management summary

This research is executed at Ceva Logistics Heerlen, a logistics service provider specialised for healthcare companies. Within medical logistics, orders must be shipped quickly, meaning that Ceva operates under short lead times. To process the orders, Ceva can schedule temporary operators per task in addition to its permanent staff on a daily basis. Despite the fact that the number of required employees only has to be communicated to the employment agency one day in advance, on average around 59% of the demand is then still uncertain. To avoid both backorders as well as workforce overcapacity, we investigate how we can align the workforce schedule to the uncertain demand as accurately as possible.

Research question

The purpose of this research is to construct a decision-support tool which is capable of generating the workforce schedule for the next business day, upon which managers can require the temporary operators at the employment agency. Operational workforce scheduling is challenging as it consists of the following steps: predicting the demand, determining the standard times for each task and finally determining the optimal number of employees per task. Therefore, the following research question has been formulated:

How can workforce scheduling of operational employees be improved at Ceva Logistics Heerlen by including short-term demand forecasting and the estimation of standard times?

Model approach

The first phase of this research analyses what information and data is available to improve the workforce scheduling method. To successfully process orders, each order follows one of six workflows, which is a predefined sequence of tasks. The workflows are related to the material handling equipment (MHE) with which the products are picked. The four MHE types are HOPT, EPT, Reach and Trolley, where trolley picks are subdivided in three subcategories, depending on the packing process.

To construct a decision-support tool that is capable of generating the workforce schedule, we follow the process as shown in Figure 1. First, we extract the historical order data of the past two years from the WMS system SAP. To use this data as input for the forecast models, we cleaned and filtered the data per MHE type.
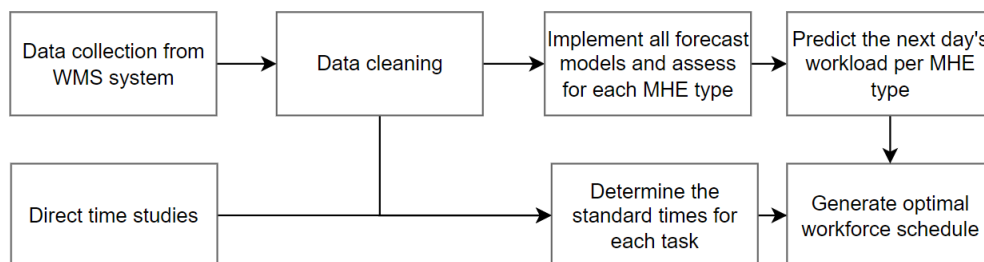

Figure 1: Process steps to create the decision-support tool capable to generate the workforce schedule

The literature study shows that within the statistical forecast models, the *Error, Trend and Seasonality* (ETS) model and the *Autoregressive Integrated Moving Average* (ARIMA) models yield good results in similar circumstances. We have implemented both models as well as a combination forecast of the ETS, ARIMA and Naïve model, for all four MHE types to predict the next day's workload. In addition, the top-down approach is used to forecast the aggregated pick lines via the ETS, ARIMA and combination models. Lastly, via the linear regression based model, we determined the linear relationship between the known pick lines and the expected total pick lines. We validated and assessed each forecast model by using cross-validation to calculate the forecast error, expressed as MAPE value, over the validation set.

As the process times of some tasks are stored in SAP, this data is used as input to determine the standard times per task. For the tasks for which no historical data is available, the standard times are determined via direct time studies.

UNIVERSITY OF TWENTE.

The goal of the workforce scheduling tool is to generate a realistic workforce schedule which is capable of aligning the workforce to the expected demand based on the estimated standard times. To create the workforce schedule, first an initial workforce schedule is generated, which consists of the absolute minimum workforce capacity required. Via the Flexible Job-Shop model (FJSP), the expected backorders and expected shift ends are determined. If there are backorders expected, an operator is added to the bottleneck task. Every iteration the FJSP model is recalculated until there are no more expected backorders. The workforce schedule that generates no backorders is selected as the solution. We selected the FJSP model as it is capable of processing all the orders according to the predefined sequence of tasks. To solve the FJSP in polynomial time, we use the *First Come First Served* and *Earliest Due Date* priority rules.

Results of the forecast and workforce schedule

Figure 2 shows the MAPE values of each forecast model per MHE type. The Trolley orders can be forecasted the most accurately, this is important as more than 94% of all orders are trolley orders. An average MAPE value of 6.6% on average is achieved, measured over a period of 55 business days. Compared to the Naïve trolley forecast Ceva currently uses, this reduces the forecast error on average by 4.7%.
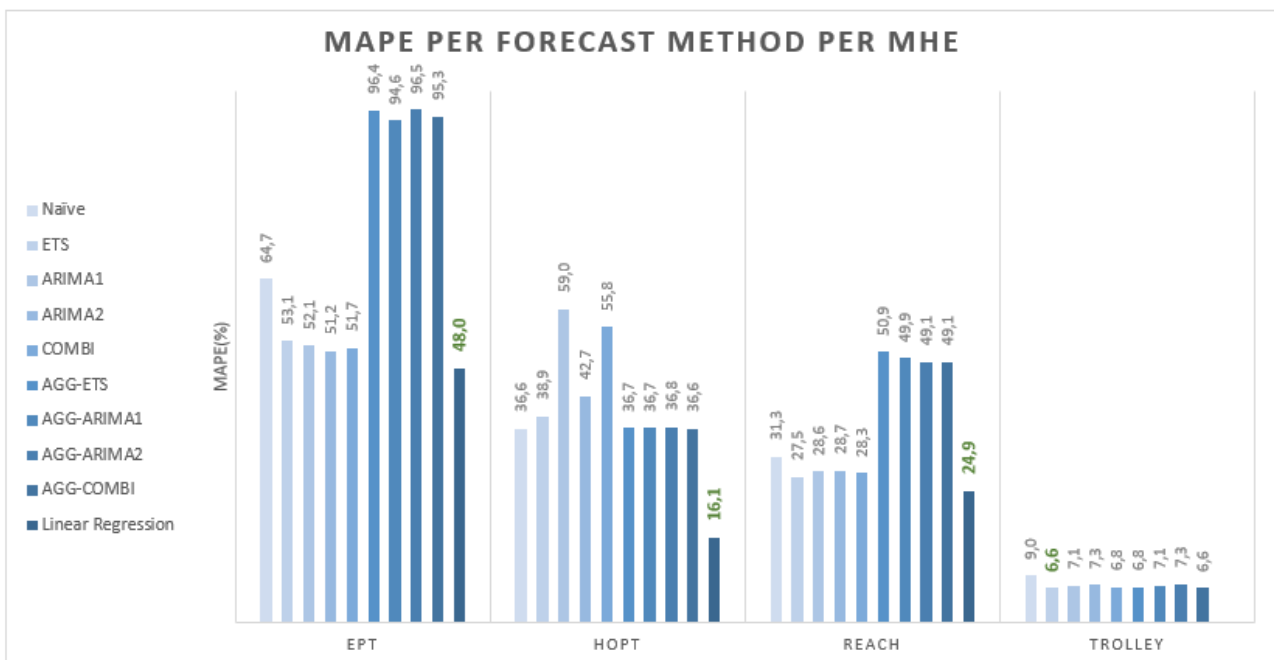

Figure 2: Forecast Error expressed in MAPE values per forecast model and per MHE type

For EPT, HOPT and REACH orders, the best accuracy is achieved by the linear regression model. The linear regression model bases the expected number of pick lines, which is the dependent variable, on the known pick lines of that respective MHE type, which represents the independent variable.

The workforce solution is intended as a decision-support tool for the manager to request the right number of temporary operators. We programmed an implementation in excel, where a macro displays the workforce schedule as Gantt Chart. The manager gains insight in the required number of FTE per task, the expected idle time and shift end per task. Based on the outcome of the workforce schedule, the manager can require the temporary operators at the employment agency.

Recommendations

We recommend building upon this research by expanding the forecast model and workforce scheduling tool to the other departments within the company. This is in line with the long-term business goal to centralise the forecasting and resource planning by creating a control room. Ultimately, it is intended that the control room will use online workforce scheduling to optimally (re)allocate the available workforce. This will contribute to the optimal deployment of staff.

Table of contents

UNIVERSITY OF TWENTE.

List of figures

List of tables

Glossary

ANN	Artificial Neural Network
EPT	Weighted Electric Pallet Truck
ES	Exponential Smoothing
ETS	Error Trend Seasonality model
FJSP	Flexible Job-Shop Schedule Problem
FTE	Full Time Equivalent
HOPT	Similar to a forklift with the difference that the operator moves up with the HOPT and piece picks the products directly from the pallet location. This is mainly used for slow mover products.
KPI	Key Performance Indicator
MHE	Material Handling Equipment, is mechanical equipment to move materials
PGI	Post Good Issue. Once the order is complete PGI is done, this moves the ownership to the customer.
Reach	A forklift that is designed for use in narrow aisles
SAP	Enterprise Resource Planning (ERP) system
SARIMA	Seasonal Autoregressive Integrated Moving Average model
SMA	Simple Moving Average
Trolley	A small movable shelving rack where the operator can store the products after he has collected them
WMS	Warehouse Management System

1 Introduction

The objective of this research is to gain insight into how a more accurate workforce scheduling tool can be beneficial for Ceva Heerlen by including a forecasting model and the standard process times per task. This quantitative research is conducted at Ceva Logistics in Heerlen, the Netherlands. In this chapter a global overview is given of the problem statement, research approach and the scope of the research.

1.1 Ceva Logistics

Ceva Logistics offers end-to-end logistics solutions to their clients. Hence, products of their clients are kept in stock in Ceva's warehouses and once an order comes in, the products are picked, packed and shipped directly to the end customer. They operate within 160 different countries around the globe and obtain a combined revenue of 7,4 billion dollars. Their warehouses offer nine million square metres of storage space in 750 locations. In total Ceva employs 78000 people worldwide.

1.1.1 Ceva Logistics Heerlen

Ceva Logistics Heerlen has specialised itself in logistic solutions for medical client companies and takes care of the total logistic process for three clients. This research focuses on the largest client as more than 75% of all the processed packages are destined for end customers of the client. Hereafter we refer to the largest client simply as client for confidentiality reasons.

As the client is responsible for the majority of the workload, the client's processes are spread over several departments. Within this research, we focus on the department responsible for shipping minimally invasive therapies products within the European Union. Figure 1.1 provides a global overview of the most common logistic process of this department.

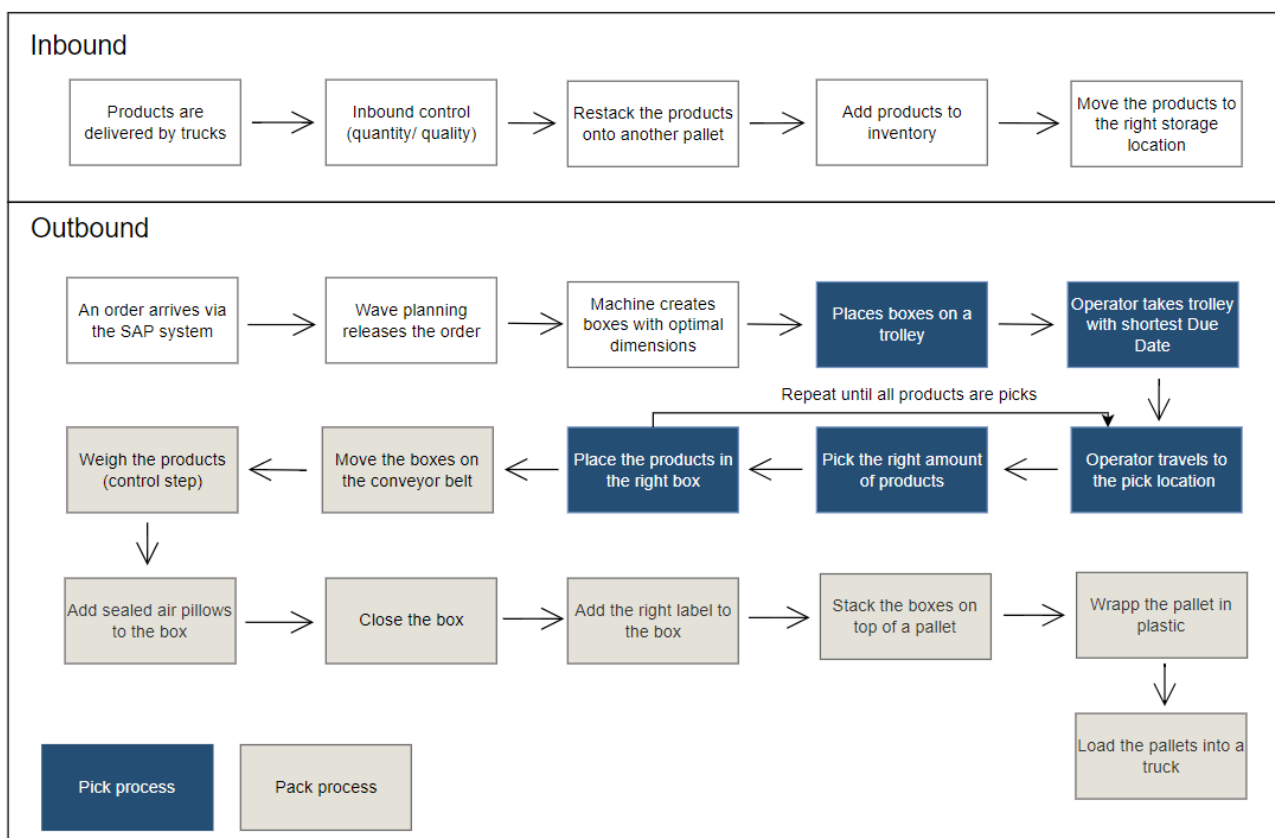

Figure 1.1: Flowchart of the global logistic process at Ceva logistic Heerlen

The products are delivered by truck and unloaded at the inbound loading docks. The inbound team checks the products in terms of quantity and quality. When multiple products are stacked on a single pallet, the decision could be made to restack some products onto another pallet. The next step is to add the products

digitally to the inventory via the warehouse management system. Based on the product type, size and quantity, the system suggests the optimal storage place after which the products are physically moved.

Once an order comes in via the client's SAP system, the outbound process starts. The wave planning department decides when to release the order to the operators based on the priority. Picking can be done by forklift, electric pallet truck or manually with a trolley or platform lift. Figure 1.1 shows the regular trolley process, for which a packsize machine creates the boxes with the optimal dimensions based on the items that will be placed in the box. After the boxes are created a trolley prepare operator places the boxes on the right trolley.

Trolley picking is performed according to the *man-to-items* system, where pickers walk to the pick locations. The operators pick the products in batches using the *sort-while-pick* method. Meaning that operators carry boxes with them on a trolley and walk through the warehouse to the products that need to be picked. The products are directly sorted by placing them in the right box. Once the boxes are filled, they are placed on a conveyor belt. The products are then automatically weighed to check if the weight matches the expected weight, this way mistakes like a missing product can be detected. Subsequently, sealed air pillows are added to the box to protect the items, after which the box is closed and the sending label is added. Lastly, the boxes are stacked on pallets and eventually loaded into trucks.

In case of bulk picks, a pallet in inventory can be shipped at once. A forklift moves the pallet from the racks to the ground after which a pallet jack is used to move the pallet directly to the shipping area. The pallet jack is equipped with a weight sensor, therefore the pallet is weighed automatically.

1.2 Problem statement

After an order comes in, the products are picked, wrapped up and sent directly to the end customer. The lead time varies per order from 90 minutes up to 48 hours. Since orders are handled within such a short timeframe, making an operational workforce schedule that matches demand is challenging.

1.2.1 Problem background

The operational workforce pool consists of permanent employees contracted by Ceva itself and temporary employees contracted via an employment agency. The permanent employees have a contract with a fixed amount of hours and are therefore assured of a constant number of hours each week. The temporary employees work on an on-call basis, therefore they are only scheduled to fill the gap between demand and the permanent workforce capacity. Ceva lets the agency know before 15:30 how many temporary employees are needed the next day. Although this construction offers Ceva a lot of flexibility to modify the workforce capacity last-minute, the exact demand is still not known at the scheduling moment as lead times are even shorter. On average 41% of the demand for the next day is known when the temporary employees are scheduled. Hence, to create a workforce schedule, a prediction has to be made about demand that will arrive after the scheduling moment.

When Ceva's site in Heerlen was built in 2017, the warehouse had an area of 33000 m^2 and 300 employees to run the daily business. Just under five years later, the site has expanded to 96000 m^2 and 650 employees. When a company grows at this speed, it is difficult to optimally adjust all the systems to the constantly changing needs of operational departments. Supervisors are in charge of requesting the appropriate number of temporary workers for each task. However, they do not possess all required knowledge to create an accurate workforce schedule which matches the expected demand. Therefore, supervisors are dependent on the current scheduling spreadsheet to determine the number of employees needed per task. The scheduling spreadsheet is a basic deterministic tool that bases the schedule on only two variables:
- A Naïve forecast: This forecast assumes that next day's demand will equal the average demand of the same week last year and multiplies it by a safety factor.
- The expected productivity: The demand that can be processed by one operator per hour.

The issue with this tool is that the workforce schedule is not accurate enough, causing a mismatch between scheduled workload and demand. This either leads to unnecessary labour costs or orders that are processed too late. As Ceva Heerlen ships medical products, it is important that products are delivered on time. Therefore, Ceva agreed upon a backorder *Key Performance Indicator* (KPI) with its client. The backorder KPI indicates the percentage of orders that are processed later than the deadline. In case the backorder KPI gets above the agreed level, Ceva has to pay a fine to its client.

As supervisors do not possess the tools or skills to proactively adapt the workforce capacity to the expected demand, they cooperate and exchange employees for several hours to balance the workload better. This is allowed within the same client's contract. Unfortunately, even by collaborating between different departments, the backorders KPI is not met. Experience shows that once Ceva is in a backorder situation, it is difficult to get out of it. This is because all orders must be processed anyway, because medical orders cannot be cancelled or delayed for a longer time period.

In order to provide better quality for the client and to avoid the penalty, it is important that the average backorder rate is reduced. The current way of scheduling in combination with the outsourcing of personnel to the various departments is no longer sufficient to achieve the agreed KPI. To determine how the agreed KPI will be achieved in the future, the problem cluster is determined in section 1.2.2.

1.2.2 Problem cluster

The action problem should indicate the discrepancy between the norm and reality [1]. Due to compelling business interest, the rates are presented as a range instead of the exact value. Therefore the action problem is defined as:
" The average of backorder rate is 6-12% measured over five months, instead it should be below 0-3%"
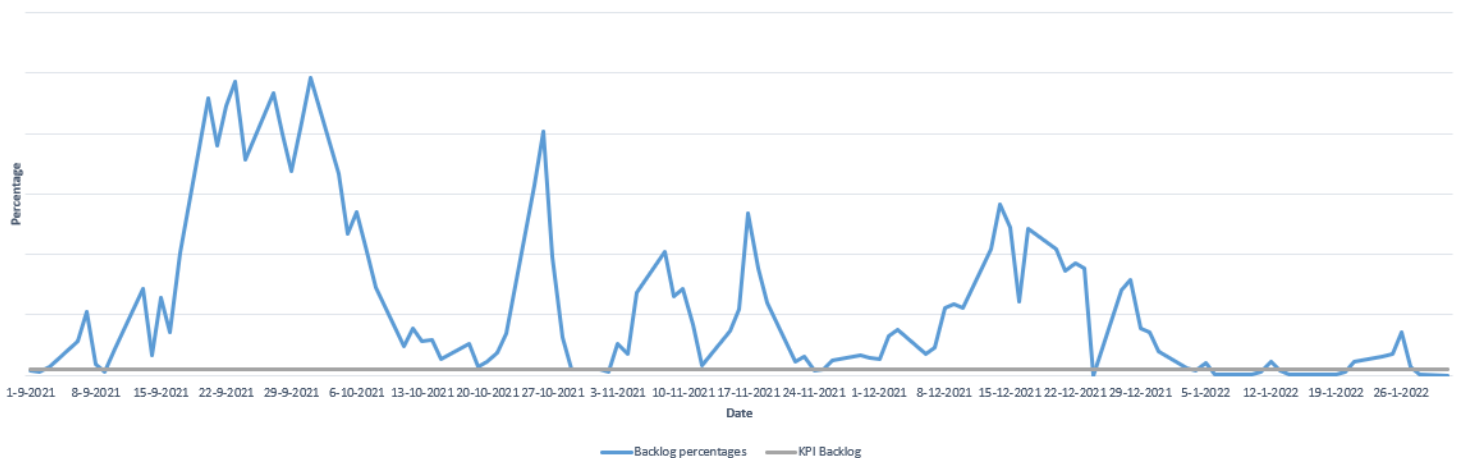
Figure 1.2: Backorder rate over time, measured over five months

Backorders are caused by the mismatch in scheduled workforce and actual demand. Figure 1.2 depicts the backorders rate of each day from the last five months in blue and the backorder KPI in grey. From the figure we observe that the days in which the KPI is met compared to the number of days at which the KPI is not met is alarming. Over these five months, only once in every four days is the backorder KPI met.

In case of a backlog situation, it is difficult to get out of it since the backorders have to be processed on top of regular demand. In addition, it is stressful for the operation to be in a backlog situation as the client receives complaints from their end-customers. This results in client requests to prioritise orders, which in turns leads to a decrease in productivity. This causes the backorder rate to increase for multiple days before the percentage drops again. Hence, this situation is therefore very undesirable. In addition, when the backorder rate is higher than 0-3%, Ceva will receive penalties since the backorder rate is one of the KPI's in the Ceva's contract with their client.

To prevent backorders, the core problems are determined via the problem cluster shown in Figure 1.3.
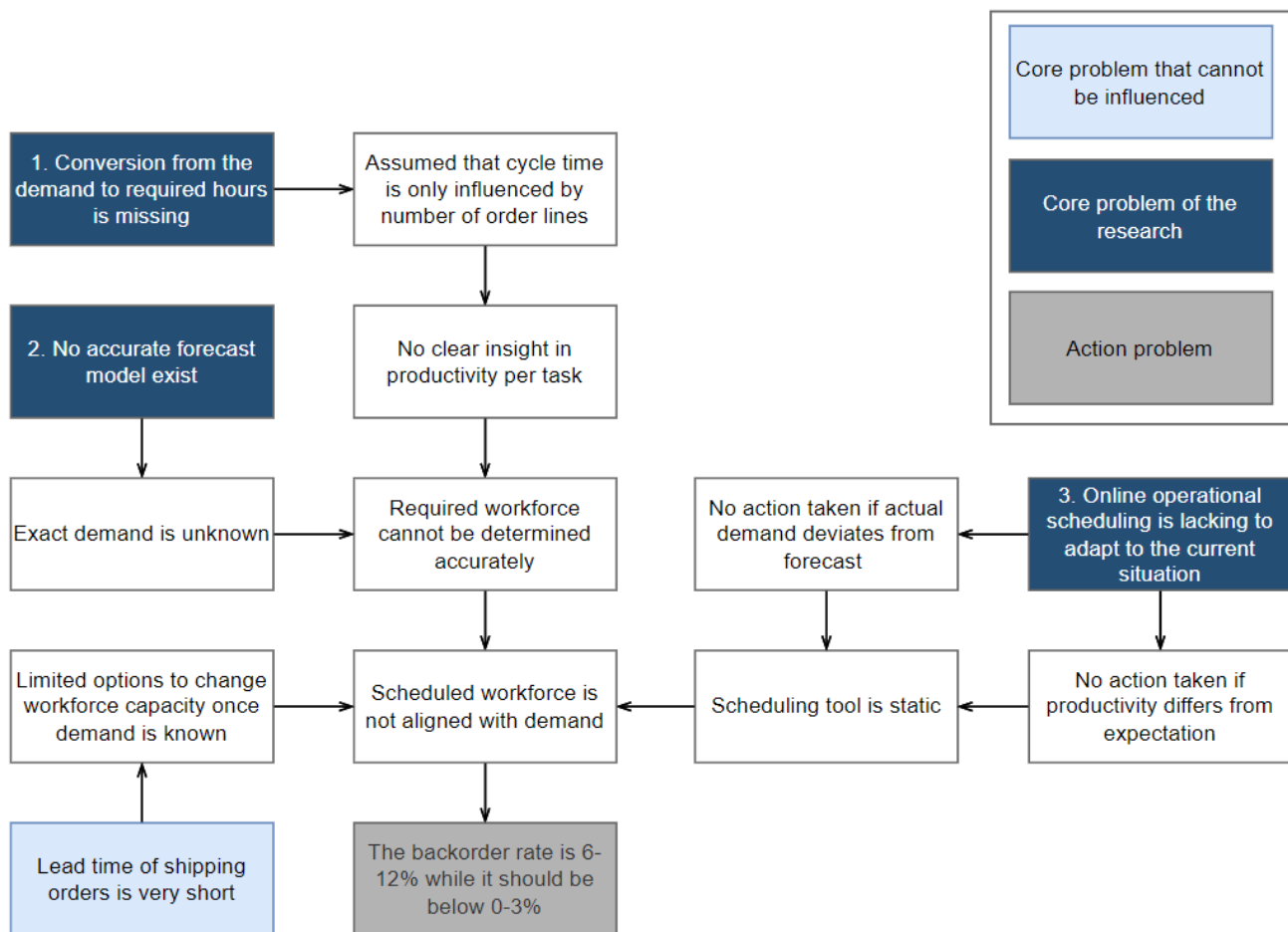
Figure 1.3: Problem Cluster

One core problem that is defined is the short lead time demand. Naturally if the lead time could be extended, scheduling the workforces would be simplified since a forecast would no longer be needed. However, to be competitive within the logistic sector, small lead times are a requirement to be competitive. To change the lead times, contract changes are required which often have financial consequences. Hence, this option is not desirable and, within the scope of this research, the lead times are assumed to be fixed.

The three defined core problems that will investigated within this research are:

1. *The translation from the demand to the required standard hours is missing.* Currently, one fixed productivity is assumed for all the pickers. While in reality, productivity depends on the type of pick. As an example, it cannot be assumed that a trolley operator picks exactly as many pick lines per hour as a forklift operator. Additionally, the workload per pick line depends on the number of products to be picked and the location where the pick must take place. Hence, Ceva needs more insight into how productivity is influenced so that the required workforce can be determined more precisely per task.

2. *No accurate forecast model.* In the current situation, the supervisor has to enter the expected pick lines into the scheduling tool. Since there is no forecasting model available, the supervisor bases his prediction on the weekly average of realised pick lines last year manually and multiplies it by a safety factor. This is a Naïve and basic way to forecast demand and it rarely leads to an accurate forecast. In conclusion, there is a need for an accurate forecasting model that guides the supervisor to select the right values as input parameters for the scheduling tool.

3. *Online scheduling is lacking in the current situation.* Orders arrive at Ceva sporadically during the day with a short lead time. Because of that, the entire order population is not known when the schedule is created. Nevertheless, the current scheduling spreadsheet is used once a day to schedule the workforce for the next day. Therefore, deviations between the scheduled scenario and the reality are noticed very late. Online operational scheduling can be used to monitor the process and react to unforeseen or unanticipated events [2]. By using online scheduling, the supervisor can accurately monitor the process per operator skill. Gaining insight will help the supervisors to respond fast by shifting capacity between the skills and thereby reducing the risks of backorders. The same applies to the arrival of orders. If the number of incoming orders deviates significantly from the forecast, the supervisor must obtain insight into this as soon as possible. This will give the supervisor the opportunities to take the right measures so that the orders are sent on time.

Due to the time restriction, the online scheduling tool will not be implemented within this research. However, the workforce scheduling tool is built in such a way that online scheduling can be easily added in the future.

1.3 Research approach

1.3.1 Statement of research goal

Conducting this research provides several benefits to the company. First of all, Ceva receives an overview of the entire process and the associated standard times per task. This makes it easier to optimise the process in the future because the bottlenecks can easily be exposed. Furthermore, the workforce scheduling is more accurate as the number of required operators can be better estimated. This leads to less waste during quiet times and less backorders during busy times.

This research also contributes to the realisation of a long-term business goal. Ceva Heerlen wants to have an operational control room halfway through the year 2023. The control room will be responsible to create an offline workforce scheduling for all departments, based on the realised productivities, the current order book and the orders that are expected to arrive. Ultimately, it is intended that the control room use online workforce scheduling to optimal (re)allocate the available workforce. This will contribute to optimal deployment of the staff. The results of this research, the forecast model, the scheduling tool and the insights into the processes and standard times are very important pillars that form the foundation of the future control room.

1.3.2 Research questions and approach

The workforce schedule of the operational staff needs to be determined one day in advance. At that moment, the number of same-day shipments that will come in tomorrow are yet unknown. Hence, a forecast is a requirement to create an accurate workforce schedule. Additionally, detailed insight into the standard time per task is lacking, which complicates the conversion from expected pick lines to the workforce capacity required. Lastly, scheduling only once a day within such a fast-changing environment is a missed opportunity.

In order to solve all three identified core problems, the aim of this research is threefold. First of all, knowledge needs to be gained about how forecasting could be used for the benefit of Ceva Heerlen. Additionally, accurate standard times should be determined. By using the forecast and the standard times, an accurate scheduling tool can be made. Lastly, the workforce schedule should be implemented in such a way that online scheduling can be easily added to support the control room in the future. Therefore, the following main research question has been formulated:

How can workforce scheduling of operational employees be improved at Ceva Logistics Heerlen by including short-term demand forecasting and the estimation of standard times?

To answer the main research question, we formulate sub-research questions which form the common thread through chapters 2 to 6.

Chapter 2: Current situation

To answer the main research question, the current process and opportunities for improvement first need to be mapped out. The following sub-research questions are formulated to gain insight and are elaborated in chapter 2. The overview of the process is based on personal observations and meetings with experts. In addition, research is conducted about what data is already available and what data still needs to be determined manually.

1.1 How are the newly arriving orders converted into clear work instructions?
1.2 What does the current workforce scheduling process look like?
1.3 What data is available for the forecasting model?
1.4 Which different workflow processes currently exist within Ceva Heerlen?
1.5 How could the result of the schedule and forecast model be implemented within the current process?
1.6 What are the constraints and the objective function of the schedule model?

Chapter 3: Literature review

To improve the current situation, a literature research needs to be conducted on which models can be beneficial for Ceva. The following sub-research questions are answered in chapter 3. Several models are discussed, subsequently these advantages and disadvantages are compared and evaluated from the perspective of Ceva.

2.1 Which forecast models are available within the literature or used in similar environments?
2.2 How can the performance of the forecasting model be assessed?
2.3 Which scheduling models are available within the literature or used in similar environments?
2.4 How can the standard time of a task be determined?
2.5 How can online scheduling be implemented within this model to adjust the schedule based on real-time data?

Chapter 4: Model description

Subsequently, the selection of the forecast model is described to provide input for the scheduling model. Furthermore, the solution approach of the workforce scheduling tool, that takes into account the standard times per process step, is described. The sub-research questions, that are described below, will contribute to gaining the knowledge needed to create the forecast model and workforce scheduling tool. The answers to these questions are elaborated in chapter 4.

3.1 How should the data be prepared before it can be used as input data for the model?
3.2 What causes of deviations in demand can be distinguished from the available data?
3.3 How can the forecast models described in literature be modified to support Ceva in the best possible way?
3.4 How can the output of the forecast model be used as input for the workforce schedule model?
3.5 How should the workforce schedule model be implemented to fit the requirements of Ceva in the best possible way?

Chapter 5: Model results and validation

After both models are implemented the performance needs to be measured. In addition the results needs to be validated. To measure and validate the forecast and workforce scheduling model, the following sub-research questions are defined. The answers to these questions are elaborated in chapter 5.

4.1 Which validation model is selected to assess and validate the forecasting model?
4.2 How well does the forecast model perform?

4.3 How is the scheduling tool validated?

Chapter 6: Implementation

Lastly, research is conducted on the changes that are required within Ceva to fully implement both models in the daily operation. Chapter 6 answers the following sub-research question:

5.1 How can the implementation roadmap be optimally designed?

Finally Chapter 7 provides the conclusion, recommendations and limitations and scientific contribution of this research and the suggestions for further research.

1.4 Research scope

The aim of this graduate project is to create a scheduling and forecasting tool. Limited by the time that is available for this project, the choice was made to focus on the outbound processes. The forecast and scheduling tools are created for Ceva's client for order within the European Union. However, the goal is to expand the tools in the future for export shipments, other clients as well as the inbound process. Therefore, it is important that both models be written in a generic way to be able to include future extensions. The time horizon of the workforce schedule is one day. Historic demand is known for a time period of two years. The workforce scheduling of the supportive departments is out of the scope of this research. Additionally, changing the contract agreements regarding the lead time is out of the scope of this research.

1.5 Assumptions

- Usually, Ceva Heerlen only operates on business days. In backlog situations, the decision is often made to work on Saturdays and if required even Sundays. The number of working hours on such days strongly depends on the willingness of operators. Since this situation is highly undesirable, it is assumed that working during weekends is not an option.
- An operator is hired for a specific client contract. Under normal circumstances, exchanging workforce between different contracts is not allowed. However, in very rare situations it is still done, when it is in the interest of Ceva. As these situations happen only once or twice a year, it is assumed that exchanging workforce is outside the scope of this research.
- The waveplanners decide when to release pick lines to the operators. Within this research no influence can be made on the releasing order of the pick lines. Therefore, within the scheduling tool, it is assumed that an order is known once it is released by wave planning.
- Since this research is conducted on the operational level, we have decided to optimise the scheduling tool based on the total number of workforce operators scheduled. The salary costs per employee are not taken into consideration as this is not a key performance indicator at which supervisors control the operation.

2 Current situation

A detailed overview of all the digitals steps required to process an order is provided in section 2.1. Additionally, the current workforce scheduling and forecasting process of Ceva is outlined in sections 2.2. Section 2.3 gives detailed insight on the lead times of the departments. In section 2.4 and 2.5 the available information for the forecast and workforce scheduling models is described respectively. Lastly, section 2.6 analyses the new requirements and the objective function of the scheduling tool.

2.1 Conversion of orders to pick lines

This section answers sub-research question 1.1: *How are the newly arriving orders converted into clear work instructions?*

Before an order can be picked from the warehouse, it is important to first determine the skills needed for the pick and the location of the pick. Since only certified operators are allowed to use, for example, a forklift, the pick must be performed by an operator that possesses the necessary skill. Therefore, orders need to be divided into clear work instructions containing information about the skill that is required for the pick, the number of products, the pick- and put away locations. Within Ceva one specific work instruction is called a *pick line*, this is further elaborated in section 2.1.3. Once an order arrives from the client's SAP system, several steps need to be conducted before the order can be converted into the most efficient pick lines. First an overview of the steps is given after which they are explained in detail.

1. Select the most appropriate pick location(s)
2. Determine the Material Handling Equipment
3. The order is converted to pick lines

2.1.1 Selecting the picking locations

For each product within the order, the SAP system automatically determines the most optimal pick locations. Ceva has separated its warehouse in three different picking zones, the *forward area*, the *reserve area* and the *shelving area*. The layout of the department is schematically depicted in Figure 2.1, including the different picking zones. In addition the areas used for packing are depicted, these tasks are discussed in detail in section 2.5.1.
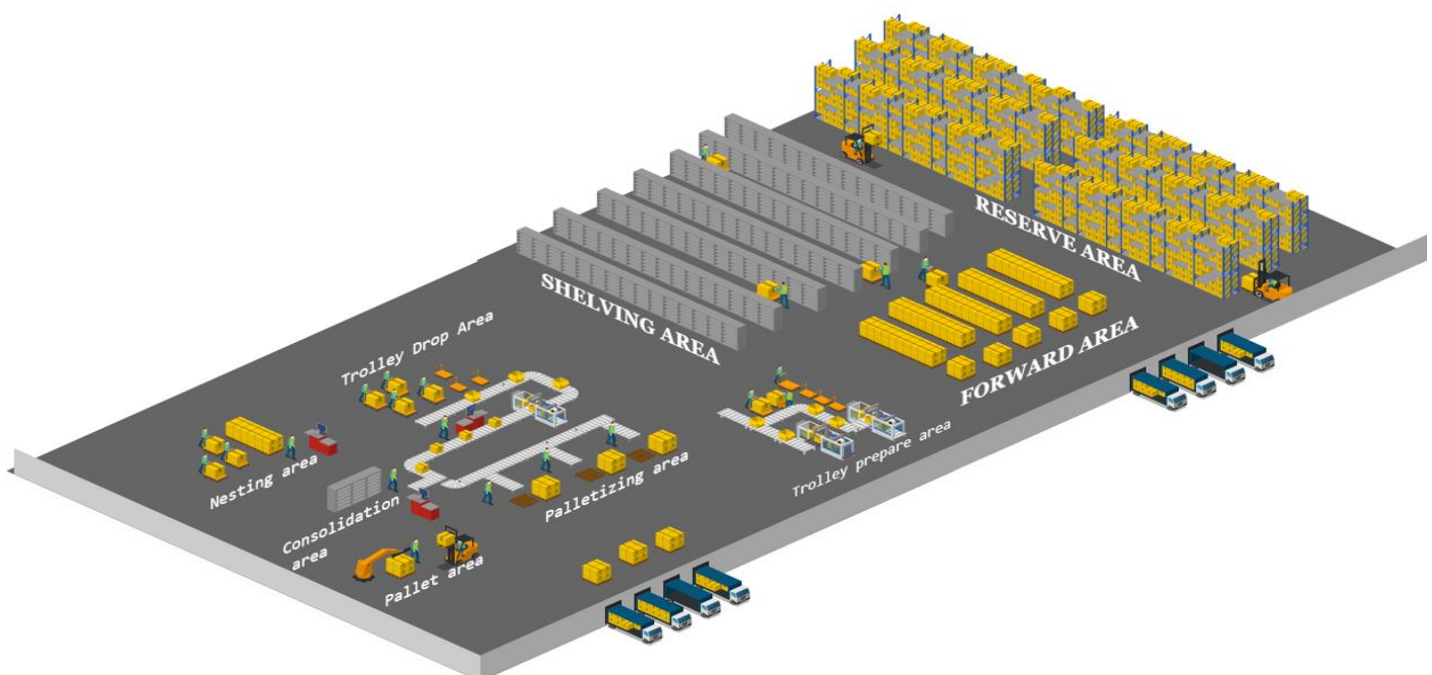
Figure 2.1: Schematic overview of storage types used at Ceva Heerlen, created in the program Icograms

In the forward area, full pallets of fast-moving products are placed directly on the ground. The goal of the forward area is to reduce the walking distance for frequently picked products, by placing these products in a compact area.

The reserve area is used to store full pallets in pallet racks. The purpose of the reserve area is twofold. The first aim of this area is to store stock which is not immediately needed, so that it does not interfere with the operational processes. Slow mover products are also directly piece picked from the reserve area.

The shelving area accommodates products that are delivered to Ceva in small quantities. Storing them on a pallet would be a waste of store space. Additionally, when products are relocated to a different location, occasionally a small batch does not fit in the new location, the 'leftovers' are placed within the shelves.

2.1.2 Determining the material handling equipment

To pick the products from these picking zones, four different Material Handling Equipment (MHE) are used, which are shown in Figure 2.2. Not every MHE is able to pick products from all the zones. For example, it is not possible for a trolley operator to pick a product at a height of three metres. Hence, the selection of the MHE is dependent on the picking location.

TROLLEY   HOPT   REACHER   EPT

Figure 2.2: Overview of material handling equipment (MHE)

To determine the most appropriate MHE per product in the order, Ceva's system first checks whether a *bulk pick* is possible or whether piece picking is required. In the case of a bulk pick, the end customer requested at least the number of products stored on a full pallet in the reserve area. In that case it is most efficient to ship the whole pallet as one unit instead of repacking all the products on the pallet into boxes and restacking them back on a pallet. A bulk pick is always carried out by a Reacher.

In case a bulk pick is not possible or there are products left to be picked after the bulk pick, piece picking is required. In case of piece picking, the needed quantity is taken from the pallet and repacked in a different packaging. In practice, this means that the operator takes the needed number of products from the pallet and places them in a box that will eventually be sent to the end customer. The first information checked by the system is whether the product that has to be picked is sterilised water. Due to the weight of this product, movements are always carried out by an EPT operator.

When the product is not sterilised water, then the SAP system searches for the most optimal pick location. Since one product can be stored in multiple locations within the warehouse simultaneously, SAP is allowed to combine different pick locations, to match the required amount. In case the products can be picked in the forward, shelving area or from a ground location in the reserve area then the pick is carried out by a trolley operator. The trolley picker picks according to the *pick-to-box* principle which means that he carries multiple end customer boxes with him on the trolley and places the product directly in the right box. An example of such a trolley is shown in Figure 2.2.

When the product is not available in the forward and shelving area, the system checks if it is allowed to pick the required amount of products directly from the reserve area. If that is the case, then the selection of the

MHE is dependent on the level on which the pallet is located within the reserve area. Piece picking from ground location pallets can be done by a trolley operator, whereas a HOPT operator is required for picks on higher levels. If piece picking is not allowed for that specific product then a replenishment order is created for a reach operator, to restock either the forward area or a ground location within the reserve area. The decision structure, as described within this section, is repeated for each ordered product.

2.1.3   Translation from order to pick lines

As described, the system must convert the order into pick lines so that the operator has the required knowledge to perform the pick. After the location and MHE are determined, all required data is available to create the pick lines. A pick line can be best imagined as a clear work instruction, containing all the information that a trolley operator needs to perform the pick. An example of a pick line is shown in Table 2.1. The 'type' column in the table represents the MHE, this is an example of a trolley pick.

Table 2.1: Example of the applicable information of a pick line

PRODUCT NUMBER	TYPE	SOURCE BIN	DESTINATION BIN	BATCH	QUANTITY
2088452038180	GP2	B1238BD	8988714092	D2H0657Y	2

A trolley operator starts his route by taking a trolley filled with empty boxes. A scanner guides the operator to the next pick, by showing the source bin code on this display. The first letter of the source bin code shows in which area the operator must perform the pick. Whereas B represents the reserve area, C the forward area and D the shelving area. The source bin B1238BD means that the product is located at the twelfth aisle in the forward area within the 38th rack at the location BD. Once the operator arrives at the right location, he compares the product and batch number on the scanner with the barcodes on the product itself. If the numbers match, the product needs to be scanned to confirm that the right product is picked. The destination bin represents the box on the trolley in which the product must be placed. The number of products that the operator needs to pick is shown on the scanner.

For every unique combination of type, destination box and pick location a new pick line is created. Hence, if the operator needs to pick five of the same products from the same location which must be divided over two different destination boxes, it would result in two order pick lines.
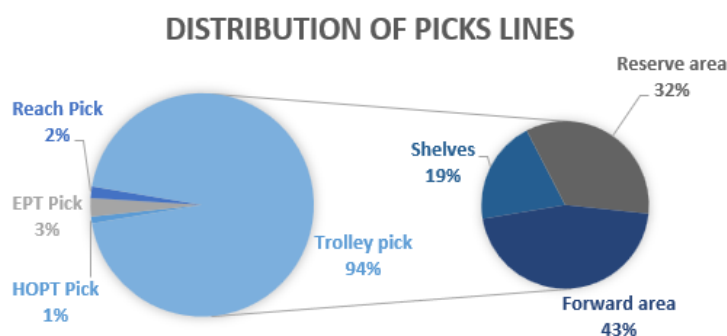
Figure 2.3: Pick Line distribution based on five months of data

Figure 2.3 depicts how the pick lines are distributed over the different MHE types. From this it immediately becomes clear that Trolley operators process 94% of all the pick lines. On the right-hand side in Figure 2.3 the distribution of Trolley pick locations is shown, as trolley pick can be performed in the forward area, the shelving area and ground locations in the reserve area. In total, 40% of all picks can be directly executed from the forward area. This is desired as the Forward area is closest to trolley pick and trolley drop locations and therefore minimises the walking distance.

2.2 Current workforce scheduling process

This section answers the sub-research question 1.2: *What does the current workforce scheduling process look like?*

The workforce schedule, which is analysed during this research, is made daily by the supervisors. Since Ceva Logistic Heerlen works within two shifts of eight hours each, there are also two supervisors who each have their own team of permanent employees. The goal of the current workforce scheduling tool is to provide the correct number of FTEs needed per shift and per task to the employment agency. In order to gain this information, the supervisors have to take two steps:

1. Determine the required number of FTE needed per task for both shifts.
2. Assign their permanent employees to a specific task and shift, so that it can be calculated how many temporary workers are needed to achieve the required capacity per task for both shifts.

The considered department has 18 different skills, for both the picking process and the packing process exactly nine. An overview of the skills and an explanation of the activities is described in Appendix A. To determine the required number of FTEs, supervisors make a division between volume dependent and volume independent tasks. Volume independent tasks require a fixed number of operators per shift, independent of the expected demand. For example, a technical operator needs to be present at all times in case a machine malfunctions.

Figure 2.4 gives an overview of the workforce scheduling per task. The number of volume independent FTEs per shift is shown in blue. This accounts for 67,5% of the total scheduled operational workforce. Based on this information, the supervisor would schedule 56 FTE regardless of the expected number of pick lines. To put this in perspective, since the department has contracted only 25 permanent employees, then at least 31 temporary employees must work each day. At first glance, it appears that some of the tasks classified as volume independent are actually volume dependent. This will be discussed in more detail in section 2.5.

Figure 2.4: The flexible and fixed FTEs needed per task per shift

In grey, the volume dependent FTEs are shown. For parcel packing, EPT and Reach picking a limited number of FTEs can be scheduled additionally during busy periods. In the spreadsheet, the optimal number of FTEs is calculated using a basic logic. To give an example, the number of EPT operators needed per shift is determined as follows:

$$IF\ (expected\ order\ lines > 8000)\ THEN\ 2\ fte\ ELSE\ 1\ fte$$

To determine the optimal number of Trolley operators, a more extended logic is used. As described earlier, the Trolley operators have to process most of the pick lines and as a result they have the highest share in the number of FTE's as shown in Figure 2.4. The required number of trolley operators is calculated based on the expected pick lines of the next business day and the expected picking rate, therefore trolley picking

does not have a minimum nor a maximum number of FTE. On average, ten trolley operators are needed per shift, which is shown in Figure 2.4.

The supervisors do not have extensive knowledge of forecasting, therefore an adjusted seasonal Naïve forecast is used to estimate the number of next day pick lines. The supervisor assumes a constant number of pick lines over the whole week, based on the average pick lines measured over the same week last year multiplied with a safety factor. In addition, the supervisor must enter how many lines per hour the average operator can collect. For this he must rely much more on his intuition, because if for a time more lines have been picked than is feasible according to the planning, then this number may be adjusted upwards and vice versa.

In conclusion, the scheduling freedom within this tool is very limited since the only task that does not have any restrictions is the trolley picking task. In addition, there is reasonable doubt whether the logic rules used really leads to an optimal schedule.

2.3 Insight lead times

When the supervisor makes the workforce schedule for the next business day, around 15:00, currently it is assumed that all of the workload of the next day is completely unknown. Therefore, he forecasts the workload exclusively based on last year's demand and multiplies it by a safety factor. However, in reality, part of the next day's workload is well-known at this point in time. It is of course a pity to predict workload that is already known, because it will increase the deviation of the forecast. To include the known pick lines in the workforce schedule, insight needs to be gained into which part of the workload can be planned deterministically and which part has to be stochastically planned. In order to gain an overview, a data analysis is performed.

For five months of historical data it was calculated for each pick line how much time was available between the moment of order creation and the due date of the order. This time window is referred to as the lead time. Over this time period, on average 45% of the pick lines have to be processed on the same day as that they are created. An additional 51% of the pick lines have to be processed the next business day and only 4% of the pick lines are known at least two business days before the due date. To explain in detail how the pick lines are divided into known pick lines and unknown pick lines, the lead times are divided into four categories. This is explained in detail based on two consecutive regular days in January 2022 as shown in Figure 2.5.

In the example in Figure 2.5 we only focus on the pick lines that have a due date of the 7th of January. As described, the large majority of these pick lines either arrive on the 6th of January or on the 7th of January. In addition, a small fraction is even known before January the 6th. The percentages shown in Figure 2.4 indicate what percentage of the pick lines, with a due date of 07-01-2022, is known at the corresponding time. Hence, in this example, between 4PM and 5PM on the 7th of January, the last 4% of the pick lines arrived. These orders have to be processed on that same day.

CUMULATIVE PICK LINES ARRIVAL
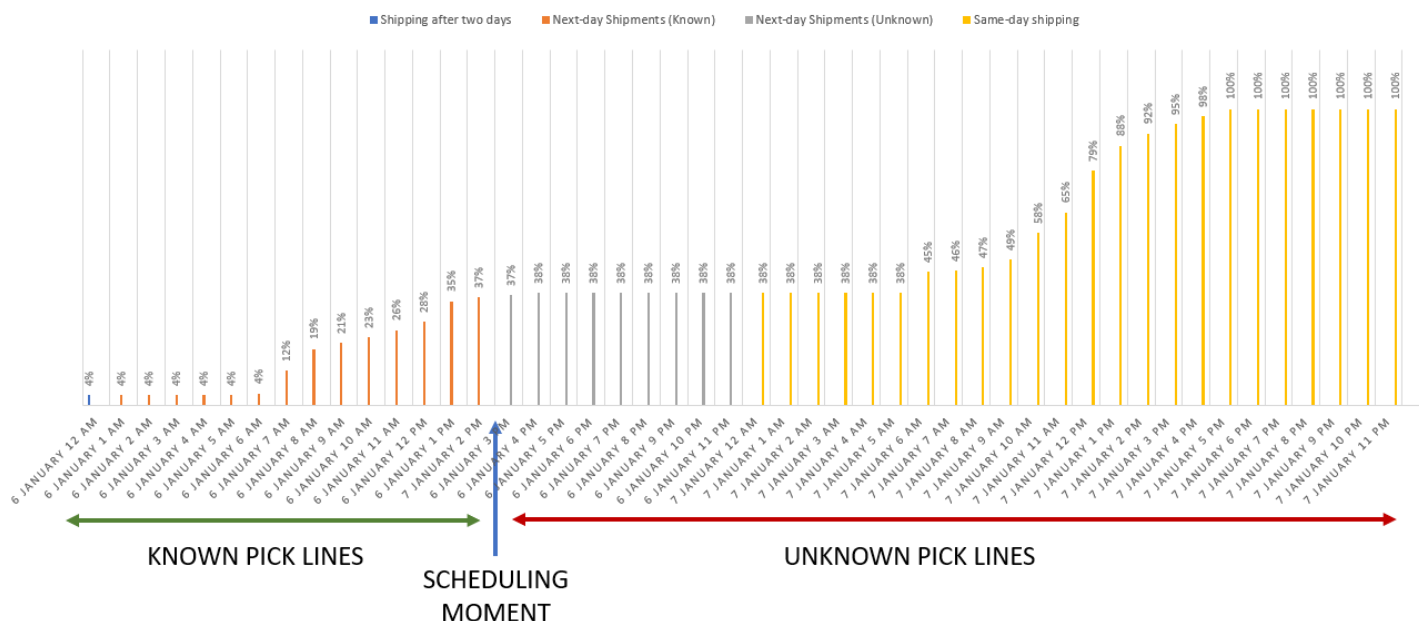
Figure 2.5: Cumulative pick lines arrival of a regular day in January 2022

The first lead time category is *Shipping after two days.* These pick lines are known at least two business days before the due date. As the workforce schedule is made one day in advance, these orders are always *known* at the moment of scheduling. These pick lines are shown in blue at the left corner Figure 2.5, where they account for 4% of the total pick lines.

Before the scheduling moment on the 6th of January, an additional 33% of the pick lines arrive. As these pick lines arrive before the scheduling moment, these are considered as *known*. This results in a total known percentage of pick lines of 37%. These pick lines fall within the category *Next-Day Shipments Known*, and are depicted in orange in Figure 2.5. We decided to set the boundary between known and unknown orders at 15:00, so that supervisors have 30 minutes to run the new scheduling tool and share the results with the employment agency.

All the pick lines that arrive on the 6th of January after the scheduling moment are considered as *unknown* pick lines and therefore fall within the category *Next-Day Shipments Unknown*. These pick lines are depicted in grey in Figure 2.5. For this particular example, only 1% of the pick lines arrive on the 6th of January after 15:00. However, it should be mentioned that this varies a lot from day to day, but on average 74% of the Next-Day Shipment pick lines arrive at the scheduling moment. Meaning that on average only 26% of the Next-Day Shipments is unknown.

The last lead time category is *Same-Day Shipments.* These are the pick lines that have to be processed on the same day as upon which they are created. As the workforce schedule is made one day in advance, these orders are always *unknown* at the moment of scheduling. These pick lines are shown in yellow at the right corner in Figure 2.5, where they account for 68% of the total pick lines.

Thus measured over five months of historical data, on average 74% of the 'Next-Day Shipment' pick lines and 100% of the 'Shipping after two days' pick lines are known when the workforce schedule is made. This means that on average $74\% * 51\% + 4\% = 41\%$ of the total orders that have to be processed the next business day is known at the latest at 15:00 today. Hence 41% of the orders can be scheduled deterministically while the remaining 58% of the orders has to be forecasted, on average. Please note that this is an average. The percentage of known pick lines can differ greatly per day. To be able to generate a good estimate of the number of unknown orders, the next section maps out which input data is available for the forecast model.

2.4 Available input forecast method

This section answers the sub-research question 1.3: *What data is available for the forecasting model?*

As mentioned in section 2.2, one input of the scheduling tool is the expected number of pick lines. In the current workforce scheduling tool, only the total expected pick lines are entered. However, since we are going to determine the standard times per task, it is important to predict the number of pick lines per MHE as accurately as possible. Therefore, Figure 2.6 shows an overview of the available data that may be relevant as input data for the forecast model and their mutual relationship.
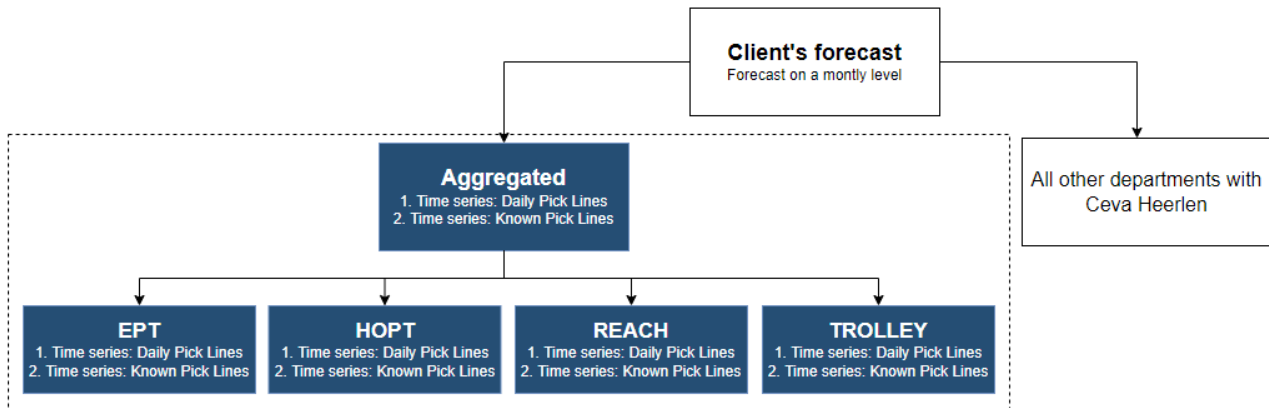
Figure 2.6: Hierarchical data that is applicable for forecasting

The available input data which is depicted in the blue rectangles, can be extracted from the client's SAP *Warehouse Management System* (WMS). This system keeps track of all the activities within the warehouse and stores, among other things, all available information of each order. This data is stored for two years and can be transformed to a time series, which is a set of observations, each one being recorded at a specific time [3]. These observations can be measured both in discrete or continuous time. The aim of forecasting time series data is to estimate how the sequence of observations will proceed into the future [4]. Therefore it is only relevant to determine the number of expected pick lines per day, the input data can be transformed to a discrete time series by aggregating the data per delivery date.

The first type of available input is the *Daily Pick Lines* time series. This is easiest to visualise as a two column table. The first column contains all dates in the past two years and the second column shows the number of pick lines that have this specific date as the due date. The Daily Pick Lines time series can be generated on a department level, meaning that it counts all the pick lines that have to be processed by the evaluated department. This is shown in Figure 2.6 as aggregated data. This data can be used as input for the forecast model, but it will estimate the expected total pick lines for the whole department. Since the result of the forecasting model should be the number of expected pick lines per MHE, the total number of lines still has to be divided over the different MHE types. This is possible because the time series data is hierarchical. Meaning that adding the time series of the individual MHE's leads to the 'aggregated' time series. By adding filters on the historical data it is also possible to determine the Daily Pick Lines time series at MHE level. Forecasting these MHE Daily Pick Lines time series individually will directly lead to the expected pick lines per MHE.

The second type of input data that is available in the WMS system is the *Known Pick Lines* time series. Pick lines are classified as Known if they are created before 15:00 and do not have to be processed today. As explained in section 2.3, on average 41% of the pick lines are known at 15:00. As with the *Daily Pick Lines* time series, the number of Known Pick Lines can be counted for each day of the last two years. The *Known Pick Lines* time series can also be determined for the aggregated level as well as for the individual MHE level. The number of known pick lines is mainly interesting in combination with the daily pick lines. It is investigated whether the number of known lines is a direct indication of the expected daily pick lines.

Now we focus on the client's forecast, depicted in the white rectangles in Figure 2.6. This forecast is the aggregated pick line over all the departments within the client's contract, which are four departments in total. In addition, this forecast is provided on a monthly level. Before this forecast can be used, the forecast must be split over the different departments and split over the days in a month. As only one department is looked at during this research, the other three departments fall outside the scope. To split the forecast, the data from all individual departments must be taken into consideration in order to achieve an accurate forecast. Hence, this is not in line with this research's scope. In addition, there are doubts within Ceva about the quality of the forecast, because there are no consequences for the client if the forecast deviates significantly. Therefore, we have decided to not include the client's forecast as input for the expected pick line forecast.

To summarise, the time series depicted in blue in Figure 2.6 are analysed to determine whether this information will contribute to an accurate forecast. As mentioned, it will also be analysed if forecasting the aggregated level or directly the individual MHEs will lead to the best result. In addition, it is examined which time series leads to the best results, the daily pick lines, the known pick lines or both. The result of the forecast is the number of expected pick lines for the next workday per MHE, hence the expected number of pick lines for EPT, HOPT, Reach and Trolley operators. This breakdown per MHE is needed to determine with the help of the scheduling tool how many employees with a certain skill are needed.

2.5 Input information available for the scheduling model

An order must be processed at the latest on its due date. Within logistics this date is often referred to as the planned PGI date. PGI stands for *Post Good Issue* which is the last step within the delivery process in which the ownership of the goods changes from Ceva to the end-customer. In practice, PGI means that the order must be picked, packed and ready to ship. To prevent an order from qualifying as a backorder, the order must not only be processed before or on the planned PGI date, but also before the planned time. The latest time that an order must be picked and packed is called the *cut-off time*, which can differ per order. The spread of the cut-off time is depicted in Figure 2.7. The latest shift of operators normally works until 22:00, so it is desired that the orders with a cut-off time after 22:00 are also finished by the end of the shift. If it was a busy day, then operators can decide to work overtime and still meet the cut-off time, but this is not a preferred situation.

To meet 100% of the cut-off times, while only 41% of the orders are known at the time of scheduling, it is necessary to build in high safety margins in the schedule. Additional workforce has to be scheduled just in case that the actual number pick lines exceeds the forecast, which leads to high personnel costs.

Fortunately, Ceva has an additional option to reduce the risk of backorders. As Ceva is not obligated to wait until the PGI date to process orders. In theory, orders can be processed directly after they arrive in the SAP system. As previously determined, on average 41% of the next day's orders are already known at 15:00. This means that if Ceva has finished processing all orders with today's PGI date, and the shift has not yet ended, the operators can start processing orders with tomorrow's PGI date. This is a huge advantage.
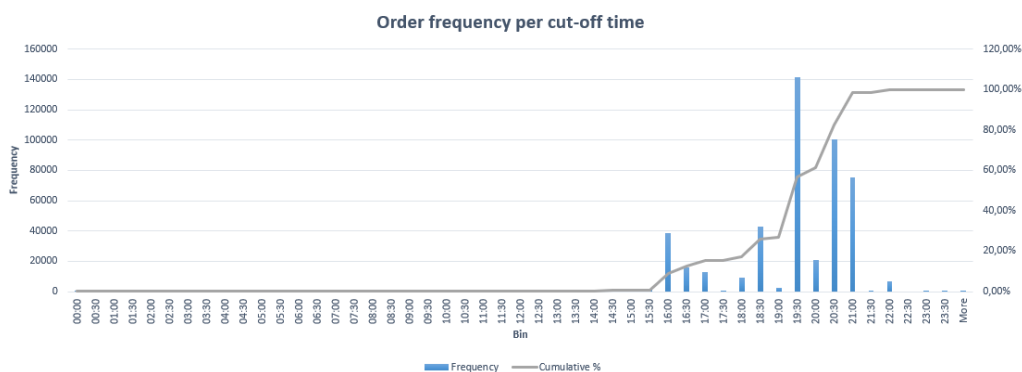
Figure 2.7: The order frequency per cut-off time, measured over five months

Suppose, when determining the workforce schedule, Ceva aims to complete all orders with today's PGI date by 18:00. This gives them the opportunity to process orders with tomorrow's PGI date from 18:00 to 22:00. Should the number of incoming orders with today's PGI date turn out to be higher than expected, they can use these four hours to process all the orders on time. The result is that fewer orders with tomorrow's PGI date can be processed. To make sure the risk of backorders is not simply postponed till tomorrow, the next day's workforce capacity can be increased. Scheduling workforce capacity according to this logic creates a safety margin to reduce the risk of backorders, without having to schedule additional operators.

2.5.1 Workflow processes

There are some tasks within the process that do not depend on the number of pick lines, these are the volume independent tasks as mentioned in section 2.2. For example, there are technical operators who ensure that malfunctions in machines are resolved. Within the current workforce scheduling tool, only four out of the eighteen tasks are marked as volume dependent tasks. However, discussion with supervisors show that is practice twelve out of the eighteen tasks are volume dependent. This significantly increase the scope of the workforce tool. To gain insight into the different tasks, sub-research question 1.4 is answered within the section: *Which different workflow processes currently exist within Ceva Heerlen?*

Through conversations with the supervisors and observing the process, six different workflows have been identified. The workflows are mainly distinguished based on MHE type, whereby trolley picks are further subdivided into three different workflows. The tasks needed to process the workflow are depicted in Figure 2.8. The first workflow is a *standard trolley pick* as described in section 1.1. The second workflow *trolley pick consolidation* requires one additional step. If an order needs to be consolidated it means that all boxes belonging to that order should be shipped on one pallet. In this case, it is allowed that there are also boxes from other orders on the same pallet. The operation solves this by removing the boxes belonging to so-called consolidation orders from the conveyor belt and temporarily putting them aside. Once all the boxes of that order are put aside, an operator places all the boxes back on the conveyor at the same time. As the boxes now move one after the other on the conveyor belt, they automatically arrive at the same pallet. Third workflow *trolley pick nesting* is similar to consolidation, but in the case of a nesting order, only the boxes of this order are allowed to be stacked on a pallet. Because no other orders are allowed on this pallet, the pallet is stacked and packed manually by the nest operator. The last three workflows are respectively the regular *HOPT pick*, *Reach pick* and *EPT pick*.

Figure 2.8: Workflows identified within the department

2.5.2 Pick line Batches

A pick line is fully processed after all steps of the relevant workflow have been completed in the right order, this must be done before the cut-off time has expired to avoid a backorder. To calculate when a pick line should be started at the latest, it is important to estimate how long it will take to carry out the workflow. Hence, for each of the twelve unique tasks as shown in Figure 2.8, the standard time must be determined.

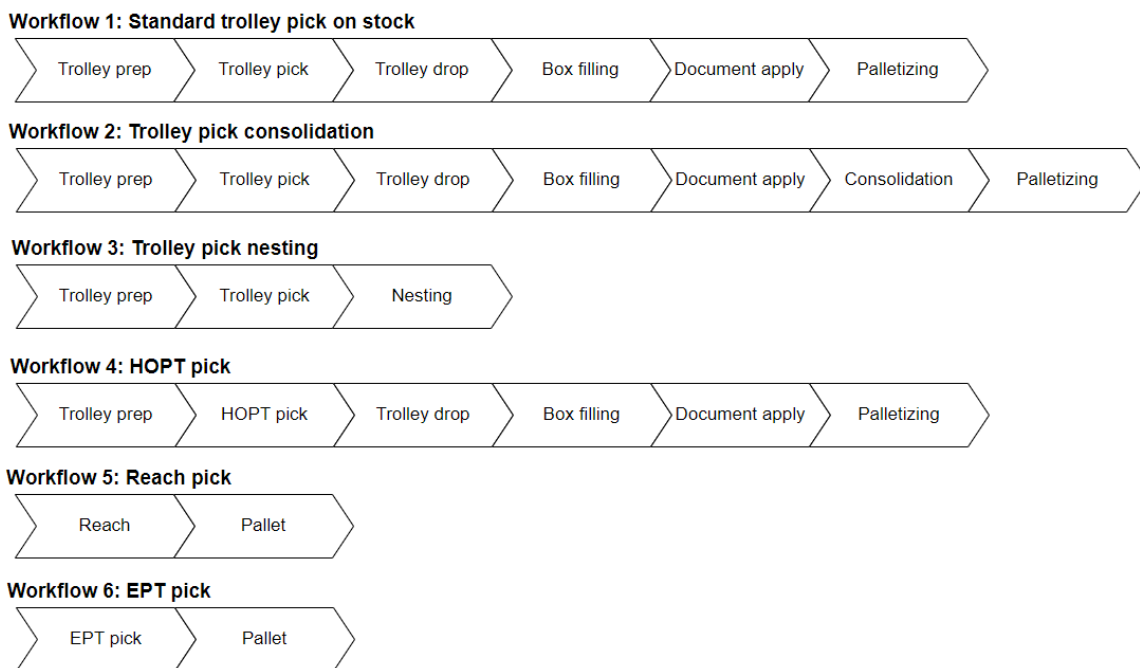In the current scheduling tool, the standard time for trolley picks is expressed in the number of pick lines per hour. After analysing the process, this does not seem to be the best unit to express the standard time. Pick lines can be linked to each other, as they may belong to the same box or trolley. Let's take the standard trolley pick as an example. A trolley prep operator prepares the trolley by placing the correct empty boxes on it so that it can be used by the trolley picker. One box can consist of one single pick line but it also can contain multiple pick lines. This does not matter for the process time, the trolley prep operator moves one box in both cases. The standard time of a trolley prep operator is therefore best expressed as the processing time per box instead of per pick line.

Each box and trolley has its own unique ID within the WMS system. For workflows 1 to 4, each pick line is linked to a box ID. All pick lines linked to the same box ID can be considered as a batch as they are processed simultaneously. Except for trolley pick, the standard time of all tasks within these workflows can be expressed in the number of boxed per hour. This not only simplifies determining the default time, but is also very useful in determining when the next step can start.

2.6 Proposed Model description and implementation

This section answers the sub-research questions 1.5 and 1.6: *How could the result of the schedule and forecast model be implemented within the current process?* and *What are the constraints and the objective function of the schedule model?*

To solve all three core problems, three different models are needed, namely a forecasting model, a workforce scheduling model and an online scheduling model. Table 2.2 provides an global overview of the three models in terms of input data, workforce capacity, result and frequency. This is then worked out in detail per model.

Table 2.2: Overview of tools created in this research to solve the core problems

	FORECAST	WORKFORCE SCHEDULE	ONLINE SCHEDULE
INPUT	1. Time series data of the last two years	1. Forecast output 2. Known pick lines 3. Standard Time per skill	1. Real-time productivity 2. Newly arrived orders 3. Workforce schedule 4. Forecast output
WORKFORCE CAPACITY	NA	Flexible	Fixed
RESULT	Expected number of pick lines per MHE	Required FTE per task	Recalculation of the optimal workforce allocation to the task
FREQUENCY	Once a day	Once a day	Regularly at a fixed time interval

Description of the forecasting model

Ceva operates within the highly competitive logistics sector, which creates a constant pressure to reduce costs while increasing the client service levels [5]. In practice this means that Ceva has to schedule as few operators as possible, while still fulfilling all cut-off times. This is important as the picking activities are responsible for 50% of the total activities within a warehouse. With that, picking is accountable for consuming the majority of the total workforce capacity within a business as well as consuming the majority

of the business's budget [6]. Combined with the small margins within the logistic sector, a precise workforce schedule is an absolute necessity. To accurately match the workforce capacity to the demand, a reliable forecast is required.

As mentioned in section 2.4, the input data of the forecast model is the historical data of a period of two years is available. as well as the high-level forecast of its client. The forecasting model has to predict the number of next day's pick lines, per MHE. Hence the forecast horizon is one business day, which is also called a *one-day-ahead* forecast. On average 58% of the pick lines are unknown at the moment that the supervisor needs to determine the required amount of flexible operators. The forecasting model needs to be calculated once every business day, so that the outcome can guide the supervisor in making this decision.

Description of the workforce schedule
The workforce scheduling process starts when the supervisor needs to determine the number of flexible employees required to process the next day's pick lines. In order to determine this, first the number of pick lines must be established. This information can partly be extracted for the SAP system, since on average 41% of the pick lines are already known, as discussed in section 2.3. The number of unknown pick lines is predicted by the forecast model. Combining this information provides insight into how many lines are likely to have to be processed the next day.

Additionally, the new workforce scheduling tool should provide the solution for the core problem: "*the conversion from the demand to the required standard hours is missing*". For this, the model must receive representative standard times per operator task as input.

Before a literature review can be conducted to map the schedule models, insight in the objective function and constraints must be obtained. Ceva has a service contract with its client in which various KPIs are recorded. A very important KPI is the backorder rate, which may not exceed the agreed value. This KPI needs to be converted into the constraints of the scheduling model. The objective of the workforce scheduling tool is to minimise the required operator hours for the department while still meeting the cut-off times. This schedule model only needs to be calculated once a day. The decision variable of this model is the number of employees per MHE type that must be scheduled.

The following hard constraints can be distinguished, hence these constraints must be met at all time:
- Only certified operators may use transport equipment to pick products. In case that there are not enough forklift certified operators available it is not possible to supplement hours with uncertified employees.
- All lines must be picked, so even orders that cannot be picked and packed before the deadline must still be processed. Therefore, unprocessed lines are added to the next day's workload.
- The number of pick lines that can be processed simultaneously is limited by the number of transport equipment that is available within Ceva. For example, if a department has three reach trucks then scheduling more than three certified Reach operators will not increase the capacity.
- The permanent employees must be scheduled for at least their contracted hours.
- The model is not allowed to schedule a negative amount of hours for any operator.
- Experts within Ceva state that a backlog situation is undesirable and it must therefore be avoided altogether. This would lead to a hard constraint of a backorder rate equal to zero, which can lead to high safety margins in terms of scheduled operating hours. That is why we will take into account in the workforce schedule that there is time between the expected last pick and the end time of the evening shift. If the demand is not higher than processed, the operators can continue working on tomorrow's orders. Therefore, no time will be wasted. In case that demand is higher than forecasted, the safety margin will prevent backorders.
- Pick lines are grouped in batches. The idea behind it is that an operator is able to pick all the pick lines within one batch in a single tour. Mainly pick lines that are stored close together in the warehouse are combined in a batch, to minimise the walking distance. Within Ceva, near optimal

order batch groups are determined by the management warehouse system. Since these order batches are sent directly to the machines that fold the boxes, it is a major change to adapt the system. It is therefore assumed that the scheduling tool cannot influence the determined order batches.

- Similar to the batch itself, the order in which a batch must be picked is determined by the WMS system. The workforce schedule is not able to influence this.
- To process an order, multiple operations must be performed in a predefined sequence. Planning the operations within a different sequence is not allowed as it does not match the reality. For example the packing operator can only start processing when the picking is ready.

Description of the online schedule

Once the supervisor has determined the number of operators needed and has forwarded the request to the employment agency, the workforce capacity is fixed. The future online scheduling tool then determines at regular intervals whether the available employees are optimally allocated to tasks. The model uses as input data, the actual number of pick lines known so far and the actual productivity per task as well as the workforce schedule and forecast model. The online schedule constantly compares the predicted scenario with the real-time data. If a significant difference is detected, it is possible that the available workforce can be used more efficiently.

The objective for the online scheduling tool is to minimise the backorders by allocating the available workforce in the most efficient way. The online scheduling tool will advise the supervisor, should the optimal workforce allocations be changed due to modified process conditions. To determine the optimal allocation, the workforce schedule model is recalculated. The fixed workforce capacity constraint is added to the workforce schedule model.

In the next chapter, an overview of forecasting models which can predict the number of expected pick lines per MHE is obtained via a literature review. This knowledge is used as input for the scheduling tool. In addition it is investigated which scheduling models could be beneficial for Ceva.

3 Literature review

In this chapter, multiple sub-research questions are answered via a literature review. Section 3.3.1 and 3.3.3 describe respectively the statistical and machine learning forecasting models that are applicable for Ceva. The chapter later explains how the model can be trained, tested and validated. Section 3.5 explains the different methods by which the standard time of a task can be determined. Section 3.6 provides an overview of the relevant scheduling models and describes how online scheduling can be implemented within a scheduling tool.

To accurately forecast and plan the needed workforce for the B-Flow operation, the process depicted in Figure 3.1 is used. The first four steps of the process fall within the forecasting scope. These are described step by step in section 3.1 to 3.4. The last step, the outcome of the forecast is used as input for the planning tool, which is described in section 3.6.
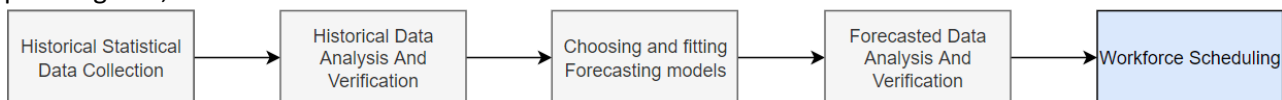
Figure 3.1: Forecasting and planning process for the quantitative approach [7]

From a methodological standpoint, forecasting and scheduling starts with the historical statistical data collection, as shown in Figure 3.1 [7]. Within this step, the required input data needs to be determined. Furthermore, within this step it is analysed how this data can be collected and with what frequency the collection should be executed.

3.1 Historical statistical data collection

For the data collection, it is important to recognise the demand drivers and perform both quantitative and qualitative data collection [7]. As Ceva logistics is a 3PL company, most of the demand drivers are unknown. For example, if the client chooses to introduce a temporary promotional price, this will most likely influence the sales quantity. However, this data is unknown for Ceva and can therefore not be taken into account.

The input data is a time series, extracted once a day from the warehouse management system as explained in section 2.6.

3.2 Data analysis and verification

The second step, data analysis and verification, answers the sub-research question: *How should the available data be prepared?*

As mentioned before, within the warehouse management systems all activities are registered. This means that in theory a lot of data can be extracted from the WMS. However, the data must be analysed before it can be used as input for the forecasting model. The data from the WMS is extracted as individual tables which need to be connected via the correct entities.

Additionally, the quality of the data needs to be examined. Luebbers, Grimmer and Jarke [8] created an overview of five important data quality dimensions in which the quality can be expressed as the following:
-    Accuracy or correctness: Does the data reflect the entities' real-world state?
-    Completeness: Does the database contain all relevant entities?
-    Consistency: Are there any contradictory facts in the database?
-    Actuality: Is the data too old for the intended purpose?
-    Relevance: Is the data relevant for the user's information needs?

By checking these dimensions step-by-step and making changes to the data where necessary, the quality is improved.

3.3 Forecasting models

Once the data is verified, it needs to be determined which forecasting model best fits the data. This section answers sub-research question 2.1: *Which forecast models are available within the literature or used in similar environments?*

To forecast the time series data, either traditional statistical methods or machine learning methods can be used. Machine learning models learn via trial and error and keep improving their performance over time, causing these methods to have longer computation times. Additionally, while machine learning models are rapidly improving, Makridakis [9] shows that traditional models are currently performing better in terms of accuracy.

Since no single method fits every problem well, the selection of a forecasting method is essential for the performance and accuracy of the forecasting model [7]. Multiple articles are studied to gain an overview of regularly used methods in similar situations. Based on Bondarev [7], Makridakis [9], Islam [10], Papaioannou [11], Züfle [12], van Gils [13] and Lalou [14], seven forecasting methods are selected to be further analysed within this research. The overview of these forecasting methods is provided in Table 3.1. The statistical methods *Holt-Winters* and *Theta* only occur only once in the analysed papers, however Theta gained the best results in a worldwide competition that compared 24 forecasting methods [9]. In this study, Theta outperforms, among others, the *ARIMA* model, which is often cited in literature. The Holt-Winters model is included in the overview, as the model is very capable to deal with seasonal patterns within time series data.

Table 3.1: Overview of the analysed forecasting methods based on [7, 9-13].

CATEGORY	METHOD	OCCURRENCE
STATISTICAL MODEL	Time series analysis: Simple Moving Average	3
	Time series analysis: Exponential Smoothing	6
	Time series analysis: Holt-Winters	1
	Time series analysis: Theta	1
	Time series analysis: ARIMA	6
MACHINE LEARNING	Regression-Based Approach	3
	Artificial Neural Network	5

Sections 3.3.1 and 3.3.3 describe the statistical models and machine learning models respectively. The statistical models, as explained in 3.3.1, can be predicted by directly using the individual MHE time series as input. In addition, it is also possible to forecast the aggregated level and then distribute it over the various MHEs, as briefly mentioned in section 2.4. Section 3.3.2 explains in detail how the aggregated forecast can be distributed across the different MHE types using the Top-Down Approach. This method can be used for any of the described statistical models.

3.3.1 Overview of the statistical models

The objective of time series statistical forecasting is to determine trends or cyclic behaviour and seasonality of pick lines based on the observations during a prior period of time, with further extension of the trends and seasonality into the future time period [7]. A trend is an increase or decrease in data over a longer period of time. A seasonal pattern means that the number of pick lines is affected by factors such as day of the week, the month or the quarter of the year. Cyclic patterns are similar to seasonal patterns, however these are not related to the calendar. An example of a cyclic pattern is the fluctuation of the stock market over time.

Determining whether a time series demonstrates seasonality, cyclic patterns or trends is also called decomposing the time series. Decomposition can either be done additively or multiplicatively [4]. In case that the seasonality appears proportional to the level of the time series, the *multiplicative decomposition* is

the most appropriate. While with *additive decomposition*, the seasonal pattern is not influenced by the level of the time series.

Simple Moving Average

The *simple moving average model* (SMA) is one of the simplest forecasting models. The models assume that the future value equals the average of m historical observations. Each historical observation y_i that is taken into consideration, is assumed equally important within this forecast. It is common to include all the available historical data. However when this leads to insufficient results, the SMA can be fine-tuned to a certain extent by adjusting the *m.* Formula 3.1 shows the calculation to predict the observation value of one period in advance $\hat{y}_{t+1}$.

$$\hat{y}_{t,t+1} = \frac{1}{m} \sum_{i=t-m}^{t} y_i \tag{3.1}$$

The advantage of this model is that it is easy to implement. A drawback of the SMA is that it reacts quite slowly to changes, as older observations are equally important. Hence, it seems to lag behind the events. SMA is most suitable to forecast on the mid- and long term.

Exponential smoothing

The *exponential smoothing* (ES) model is similar to the SMA model. The largest difference is that ES does not allocate the same weight to each historical observation which is taken into consideration. In order to let the model react faster to changes, it may be sensible to attach larger weights to more recent observations than to observations from the distant past [4]. Additionally, all the historical observations are considered in the ES model. Therefore, adjusting the accuracy of the forecast by modifying the number of observations is not possible for the ES method. Formula 3.2 represents the calculation to forecast the observation value of the next time period.

$$\hat{y}_{t,t+1} = \alpha[y_t + (1-\alpha)y_{t-1} + (1-\alpha)^2 y_{t-2} + \cdots + (1-\alpha)^m y_{t-m}] \tag{3.2}$$

α represents the smoothing parameter, which must be a value between zero and one. As the smoothing parameter decreases exponentially, observations further in the past have a smaller influence on the predicted value. The larger α, the more weight is given to recent observations.

By adding the smoothing parameter, the model can be tuned in such a way that it responds faster to changes. However, a disadvantage is that this model, like the SMA, does not take trends or seasonality into account.

Holt-Winters

The Holt model is an extended simple exponential smoothing model to allow the forecasting of data while taking into account the trend [4]. The model consists of three equations as shown in formulas 3.3 trough 3.5. The predictable value is based on the estimated level (l_t) and the extrapolation of the trend (b_t) h time periods ahead.

For each forecast, the smoothing equations for the level and the trend, equations 3.4 and 3.5, are updated. The smoothing parameters α and β^* determine, just like with the ETS method, how much weight is placed on the most recent historical observation compared to the older observations. The parameter values should lie between zero and one. The optimal parameters can be determined by minimising the least square error of the forecast.

$$\text{Forecast equation: } \hat{y}_{t,t+h} = l_t + hb_t \tag{3.3}$$
$$\text{Level equation: } l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{3.4}$$
$$\text{Trend equation: } b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \tag{3.5}$$

Holt's model takes the trend into account but does not include a seasonal pattern in the forecast yet. Therefore Holt [15] and Winters [16] extended the original model with seasonality and called the adapted model Holt-Winters. The modifications compared to the original Holt's model are depicted in black in formulas 3.6 to 3.9.

$$\text{Forecast equation: } \hat{y}_{t,t+h} = l_t + hb_t + s_{t+h-m(k+1)} \tag{3.6}$$
$$\text{Level equation: } l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{3.7}$$
$$\text{Trend equation: } b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \tag{3.8}$$
$$\text{Seasonality equation } s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)_{s_{t-m}} \tag{3.9}$$

In theory both models, Holt and Holt-Winters, can be used to predict observations multiple time periods ahead. In that case, these models assume a constant linear trend. In practice however, this method tends to over-forecast, especially on long-term forecasts [4]. To overcome this problem. Gardner & McKenzie [17] have developed a parameter that damps the trend to a flat line within the long-term. As the objective of the forecast is to predict next day's pick lines, the forecast operates within the short-term. Therefore, the damping parameter will probably not influence the one-period-ahead forecast.

ETS function

In the statistical program R, Hyndman et al [18] has written a function that is able to find the best combination of trend and seasonal components for the observed time series. The function is called ETS, which stands for Error, Trend and Seasonality. The model is capable of modelling every possible combination between the trend components [*none, additive, additive damped, multiplicative, multiplicative damped*] and the seasonal components [*none, additive and multiplication*].There are fifteen exponential smoothing methods possible, which are described in Table 3.2 [4].

Table 3.2: Classification of exponential smoothing methods [4].

TREND COMPONENT	SEASONAL COMPONENT		
	N (None)	A (Additive)	M (Multiplicative)
N (NONE)	(N,N)	(N,A)	(N,M)
A (ADDITIVE)	(A,N)	(A,A)	(A,M)
A_d(ADDITIVE DAMPED)	$(A_d,$N)	$(A_d,$A)	$(A_d,$M)
M (MULTIPLICATIVE)	(M,N)	(M,A)	(M,M)
M_d(MULTIPLICATIVE DAMPED)	$(M_d,$N)	$(M_d,$A)	$(M_d,$M)

Some of these combinations are models that have been discussed before, are:

(N,N)   Simple exponential smoothing
(A,N)   Holt's linear method
(A_d, N) Additive damped trend method
(A,A)   Additive Holt-Winters method
(A,M)   Multiplicative Holt-Winters method
(A_d, M) Holt-Winters damped method

These models do not have to be programmed separately in R because they are already included in the ETS function.

Theta

Assimakopoulos and Nikolopoulos [19] represented a new univariate forecasting method called the *Theta* model. Basically, it is a special case model of exponential smoothing with drift [20]. The objective of this model is to decompose the time series data into multiple series with individual long and short-term components. These newly generated series are identified using the model and then extrapolated separately. Subsequently, the actual forecast is calculated by combining the extrapolation of those series.

The first step of the model is to modify the local curvatures by calculating the second difference as shown in formula 3.10 [19] [20]. Solving equation 3.10 to $Y_{t,\theta}$, which is called a Theta-Line, provides formula 3.11, where a_θ and b_θ represent constant values. Thus, $Y_{t,\theta}$ is equivalent to a linear function of X_t with a linear trend added [20]. The theta model can be modified by changing the number of decomposed series and the corresponding theta values θ.

$$Y''_{t,\theta}(\theta) = \theta * X''_t, \quad where\ X''_t = X_t - 2X_{t-1} + X_{t-2} \tag{3.10}$$
$$Y_{t,\theta} = a_\theta + b_\theta(t-1) + \theta X_t \tag{3.11}$$

Assimakopoulos and Nikolopoulos [19] illustrated the model by creating two theta-lines with the Theta values of zero and two. The first theta-line ($\theta = 0$) results in a linear regression line which describes the time series through a linear trend. This theta-line contains the long-term components. The extrapolation can easily be done via the usual way for a linear trend. The second theta-line ($\theta = 2$), doubles the local curvatures magnifying the short-term behaviour. The extrapolation of this model is performed via simple exponential smoothing. To determine the predicted value, both extrapolation values are averaged.

The Theta-model resulted in a higher accuracy forecast compared to more advanced methods according to empirical studies [21]. The drawback of this model is that the time series data needs to be seasonalised, if applicable, before the model can be used. After the forecast is calculated, it needs to be re-seasonalised.

ARIMA
Together with the exponential smoothing parameters, the (S)ARIMA models are the most widely used models to forecast time series data [4]. SARIMA is an abbreviation for Seasonal Autoregressive Integrated Moving Average model. The input for this model are the ARIMA parameters (p, d, q) and the seasonal parameters $(P, D, Q)_m$. The lowercase and uppercase letters represent the same terms for respectively the ARIMA and the seasonal parameters. The letter P represents the order of the autoregressive part, the D is the degree of first differencing involved and the Q is the order of the moving average part. The parameter m is the number of periods per season. The purpose of this model is to modify the parameters in such a way that the forecast is as accurate as possible. This results in dozens of possible combinations of parameters.

As stated before, the (S)ARIMA model consists of an Autoregressive model and a moving average model, the equations are depicted respectively in formula 3.12 and 3.13. Autoregression is explained in detail in section 3.3.3. The input data for the (S)ARIMA model needs to be stationary. When the data is non-stationary, the difference of the data should be taken until it becomes stationary. In case that the original data is already stationary, then the integration will not be needed. Hence, this simplifies the model to an ARMA model. In formula 3.14, the equation of the ARMA model is given.

$$y_t = c + \sum_{i=1}^{p}(\emptyset_i y_{t-i}) + \varepsilon_t \tag{3.12}$$

$$y_t = \mu + \sum_{j=1}^{q}(\theta_j \varepsilon_{t-j}) + \varepsilon_t \tag{3.13}$$

$$y_t = c + \varepsilon_t + \sum_{i=1}^{p}(\emptyset_i y_{t-1}) + \sum_{j=1}^{q}(\theta_j \varepsilon_{t-j}) \tag{3.14}$$

Selecting the appropriate parameters can be challenging, however statistical programs like *R* and *python* have built-in packages that help selecting the appropriate parameters [4]. In addition, the *Akaike's Information criterion* (AIC) is widely used [22], the equation is given in formula 3.15. n is the number of data points , RSS represents the residual sums of squares and K represents the number of parameters. The combination of parameters that lead to the lowest AIC value is selected.

$$AIC = n ln\left(\frac{RSS}{n}\right) + 2K \tag{3.15}$$

Hyndman and Athanasopoulos [4] introduces a general approach to fit the ARIMA model to a time series data:

1. Plot the data and identify any unusual observations
2. If necessary, transform the data (using a Box-Cox transformation) to stabilise the variance
3. If the data are non-stationary, take first differences of the data until the data are stationary
4. Examine the ACF/PACF: Is an ARIMA(p,d,0) or ARIMA(0,d,q) model appropriate?
5. Try your chosen model(s), and use the AIC to search for a better model
6. Check the residuals from your chosen model by plotting the ACF of the residuals and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model
7. Once the residuals look like white noise, calculate the forecast.

3.3.2 Top-Down Approach

As mentioned in section 2.4 the time series data is available at the aggregated level and at the individual MHE level. The objective of the forecast model is to predict the number of next day's pick lines for each MHE type. If the time series of the individual MHE types is used as input for the statistical models, then the number of expected pick lines for that MHE type is directly predicted. However, if the time series is used at the aggregated level, the number of total expected lines must first be divided among the available MHE types. The latter can be done via the Top-Down approach [4]. The approach only works with strictly hierarchical aggregation structures, which is the case for the aggregated time series.

The Top-Down approach captures the average historical value of the bottom-level series $y_{j,t}$ relative to the average value of the total aggregate y_t, for each proportion p_j [4], as shown in formula 3.16. As the historical proportions do not take account changes over time, the Top-Down approach tends to produce less accurate forecasts at lower levels of the hierarchy [4].

$$p_j = \sum_{t=1}^{T} \frac{Y_{j,t}}{T} / \sum_{t=1}^{T} \frac{y_t}{T} \qquad (3.16)$$

Even though this may seem like a laborious method leading to less accurate results, the MHE type is known as soon as the order is converted into pick lines. However, this conversion does not take place in real time. Depending on the impact of this delay, we may have to rely on a Top-Down approach for all known orders that have not yet been converted to pick lines.

3.3.3 Overview of the machine learning models

Machine learning models use their experience to keep improving the performance or accuracy of predictions [23]. The most standard machine learning methods are classification, regression, ranking and clustering. Regression is the best suited for this research as it can be used to predict the real value for each item [23] which is in line with the forecast objective to predict the number of next days pick lines.

Machine learning algorithms can be classified in supervised learning and unsupervised learning. In case of supervised learning a labelled set of examples should be provided as training data, this is suitable for regression, ranking and classification problems. Unsupervised learning uses exclusively unlabelled training data, which is often used for clustering. Within this section, the Regression Based Approach and the artificial neural network is discussed in detail.

Regression based approach

Regression models are useful to explore the relationship between multiple time series [4]. A commonly used example is the relationship between gas consumption and the outside temperature. As the temperature increases, naturally the gas consumption decreases. When both time series are plotted against each other, the extent to which these variables influence each other becomes visible. It is important to visually determine the mathematical function that describes the relationship between the variables best, so whether the relation is e.g. linear or polynomial. To explain the concept of regression, it was decided to describe linear- and autoregression in more detail.

Linear regression

Linear regression is a technique which examines the relationship between a dependent variable and a specific independent variable [10]. The strength of the relationship between variables x and y is often expressed in the correlation coefficient r, which can be calculated by formula 3.17 [4].

$$r = \frac{\sum(x_t - \overline{x})(y_t - \overline{y})}{\sqrt{\sum(x_t - \overline{x})^2}\ \sqrt{\sum(y_t - \overline{y})^2}} \tag{3.17}$$

The correlation value lies between -1 and 1, where negative values indicate a negative relationship between the dependent and independent variable and positive values indicate a positive relationship [4]. Values close to zero represent a low correlation between the two variables, whereas values close to either 1 or -1 represent strong relationships. Subsequently, the future value of the dependent variable can be predicted once an overview is obtained of how significant each independent variable influenced the dependent variable in the past [10].

Based on the number of variables that significantly influence the dependent variable a distinction can be made between simple linear regression and multiple linear regression. A simple linear regression model consists of a single independent variable, whereas multiple linear regressions models consist of at least two independent variables. Both models can be modelled with formula 3.18, where $\beta_2 x_2 + \cdots + \beta_k x_k$ is equal to zero for the single linear regression model. As in every linear function, β_0 represents the intersection of the x-axis, while β_1 till β_k represent the effect of each independent variable after taking into account all the other applicable independent variables. ε_t is the deviation between the actual measurement and the expected value.

$$y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon_t \tag{3.18}$$

Autoregression

Autoregression is similar to linear regression, however it measures the linear relationship between lagged values of a single time series[4] instead of the relationship between multiple time series. Via the autocorrelation coefficient r_k, it can be determined whether the time series data follows a cyclic or seasonal pattern. The formula to calculate r_k is shown in formula 3.19, where k represents the lagged period. Hence, r_2 measures the relationship between the current time period and the value of two time periods ago.

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t - \overline{y})(y_{t-k} - \overline{y})}{\sum_{t=1}^{T}(y_t - \overline{y})^2} \tag{3.19}$$

Next, the values of r_k can be depicted in an autocorrelation function (ACF). An example of quarterly beer production is shown in . The blue dotted lines distinguish the r_k values in 'values significantly different from zero' and 'white noise'. Bars that are out of the boundaries represent a significant relationship. Smaller bars which do not exceed the boundary values are called white noise. The most significant values for the autocorrelation coefficient are r_2 and r_4. r_4 represents a positive relationship between the current beer production and the production of four quarters ago, meaning that the variation of the time series data follows a yearly seasonal pattern. In addition, r_2 shows a negative relationship between lagged data of two quarters ago. This represents a high beer consumption within the summer period and a low production within the winter period. The remainder of significant values are repetitions of r_2 and r_4.
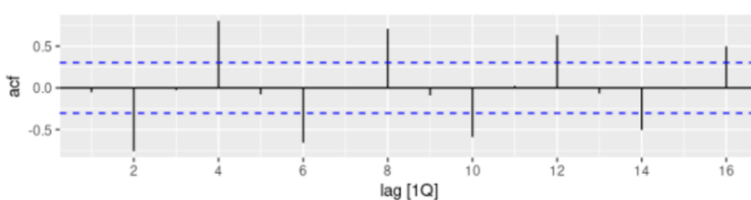

Figure 3.2: Example of an autocorrelation function of quarterly beer production [4]

Forecasting with Regression Based Approach

Regression based models can be used to predict the real value for each item. Depending on the magnitude of the error between the predicted value and the real value, a penalty is given [23].

In practice, the coefficients $\beta_0, \beta_1, \dots \beta_k$ need to be determined based on the observations from the time series data. The *least squares estimation* is a possible way to determine these coefficients effectively by minimising the sum of the squared errors [4], the equation is shown in formula 3.20.

$$\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} \left(y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \cdots - \beta_k x_{k-t} \right)^2 \tag{3.20}$$

Artificial Neural Networks

A more advanced machine learning model is the *artificial neural network* (ANN). The objective of ANNs is to simulate the intelligence of the human brain by a machine. The model tries to recognise regularities and patterns in the input data, learn via experience and convert its knowledge into a result based on its knowledge gathered so far [24]. The model is similar to the linear regression model, however the ANNs can contain nonlinearities, which are captured within an additional hidden layer as depicted in Figure 3.3.
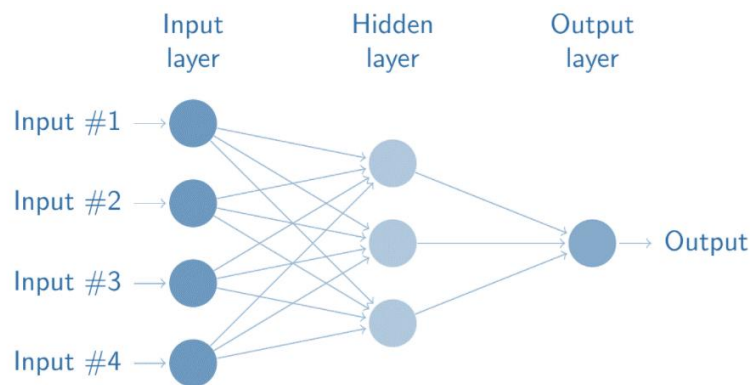
Figure 3.3: Schematic overview of a neural network [4]

The output of the model can be computed via equation 3.21 [24].

$$Y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \, g \left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i} \right) + \varepsilon_t \tag{3.21}$$

The explanation of the abbreviations used in the equations is:

Y_t	Output value	α_j, β_j	Connection weights
Y_{t-i}	Input value	ε_t	Random Shock
P	Number of input nodes	α_0, β_{0j}	Bias terms
q	Number of hidden nodes		

Activation functions are used to describe the output of a node given a (set of) inputs. Usually, the logistic sigmoid function $g(x) = \frac{1}{1+e^{-x}}$ is used as the nonlinear activation function. Next, the connection weight needs to be determined. This can be done by minimising the error function.

Time Lagged Neural Networks is a variation of the ANN model in which the input nodes are using the same time series values but with a different delay. This can be compared well with the difference between linear regression and autoregression.

ANN models can process the data to approximate a large class of functions parallel, while achieving a high degree of accuracy [24]. Additional benefit of the ANNs model is that it is able to deal with incomplete or fuzzy input data.

3.3.4 Overview of the introduced forecast models

Table 3.3 provides an overview of the performance of all mentioned forecasting models, based on the computation time (CT), the required input data and the results. Two frequently used functions in the statistical program R covers the majority of the introduced statistical models. The first function in r that is

used to forecast is the *ETS function*. As mentioned before, this function covers, among others, the exponential smoothing and Holt-Winters model. The second function is the *Arima function*. By using the moving average and autoregressive component of the model, Arima will result in an accuracy at least as good as the simple moving average. For example: in case the SMA results in the best accuracy, Arima will simply set the other components equal to match the accuracy.

The only statistical model which is not covered by the ETS and Arima function is the Theta method, however as seasonality is applicable for the input data, it is assumed that this will not lead to the best result. Hence the ETS function, Arima function, regression based analysis model and the artificial neural network models are evaluated further in Chapter 4.

In addition to linear regression based models, the neural based model was also discussed in Chapter 3 as a possible forecast model to forecast the number of pick lines. However, neural networks are mainly used when there are several independent variables that influence the dependent variable. The model can then learn what weight should be given to which variable in order to determine the dependent variable as accurately as possible. As described in detail in sector 4.1.1, only two variables are available in this research. Namely, the daily pick lines of the last two years and the number of known pick lines. The neural based network is not analysed further in the rest of the research.

Table 3.3: Overview of the described forecasting methods. SM = statistical model, ML = Machine learning

	METHOD	COMPUTE (CT)	INPUT DATA	RESULTS
SM	SMA	Very easy	Very little as the horizon can be tuned	Most suitable for mid- and long term forecasts, as it react quite slowly to changes on the short-term
	ES	Extremely easy[15]	Minimum data required[15]	Results in a slighter lower sMAPE compared to SMA [9]
	Holt-Winters	Easy	To determine seasonality, multiple cycles are required as input data	For short-term forecasting, the sMAPE equals ES. Hence, it outperforms SMA [21]
	Theta	Very easy [21], however data needs to be de-seasonalised	Similar to Holt-Winters as seasonal patterns need to be distinguished	Works very well for non-seasonal data or data with single seasonality [21]. Results in the lowest sMAPE [20]
	(S)ARIMA	Needs more CT to reach a low MAPE, compared to the other TS models.	Requires a large number of observations to determine the best fit [25]	Results in a slighter higher sMAPE compared to the Theta model [8]
ML	RBA ANN	ML models need More CT needed compared to TS	These models need a lot of input data in order determine the weight factors	Although, ML models improve rapidly, within the M3 competition they resulted in a higher sMAPE [8]

3.4 Forecasted data analysis and verification

This section answers sub-research question 2.2: *How can the performance of the forecasting model be assessed?*

In literature there exist multiple models to calculate the difference between the observed value and its forecast, which is often also called the *error (e_t)*. These calculation methods can be divided into three categories [4].

Scale dependent errors are only dependent on the error, therefore these calculations are not useful to compare different time series forecasts with each other. The two most common scale dependent error calculations are presented in equations 3.22 and 3.23. When the forecast parameters are tuned to minimise the MAE, the forecast will approach the median while minimising the RMSE leads to a forecast that approaches the mean.

$$Mean\ absolute\ error: MAE = mean(|e_t|) \tag{3.22}$$

$$Root\ mean\ square\ error: RMSE = \sqrt{mean(e_t^2)} \tag{3.23}$$

It is also possible to give the *error as a percentage* of the time series values, which makes it possible to compare different time series with each other. The most frequently used error calculation is the MAPE, which is shown in formula 3.24, y_t represents the observed data. The disadvantage of the MAPE is that too high or too low forecasts are not penalised equally severely. A forecast cannot be negative, which means that a forecast which is too low, can maximumly result in a MAPE of 100. A forecast that is too high does not have such an upper limit.

$$Mean\ absolute\ percentage\ error: MAPE = mean\left(\left|\frac{100 * e_t}{y_t}\right|\right) \tag{3.24}$$

To overcome this problem, the symmetric MAPE, also called sMAPE, is introduced. It can be calculated according to formula 3.25.

$$sMAPE = mean\left(\frac{200 * |y_t - \widehat{y_t}|}{(y_t + \widehat{y_t})}\right) \tag{3.25}$$

The last category are the *scaled errors,* which can include the accuracy comparison in the calculation. Equation 3.26 shows the calculation of the scale error q_j, which results in a value less than one if the forecast is more accurate than the simple forecast. When the simple forecast is more accurate than the evaluated forecast q_j, it results in a value above 1.

$$q_j = \frac{e_t}{\frac{1}{T - m}\sum_{t=m+1}^{T}|y_t - y_{t-m}|} \tag{3.26}$$

From q_j the *mean absolute scaled error* MASE and the *root mean square error* RMSSE, can simply be

calculated by respectively $MASE = mean(|q_j|)$ and $RMSSE = \sqrt{Mean(q_j^2)}$.

Once the accuracy of the forecast is calculated, it is common to compare it to the accuracy of a more Naïve model like the SMA or ES [4]. This gives a good impression whether the extra computation time of the more complex model outweighs the improvement in accuracy. It should be noted that a more complex model will not always outperform a simple model. In that case, it is worth considering implementing a simpler model.

Validation of the forecast
A potential consequence of selecting a forecast only based on the forecast error, is that some forecast models tend to *overfit* the data. This means that the statistical model fits the training data (almost) perfectly as shown in Figure 3.4. This is achieved when the model memorises the noise within the historical data and therefore creates a complex model. The problem with overfitting is that the model seems to be a good forecast based on the training data, however when data is used that the model has never seen it leads to quite poor results.
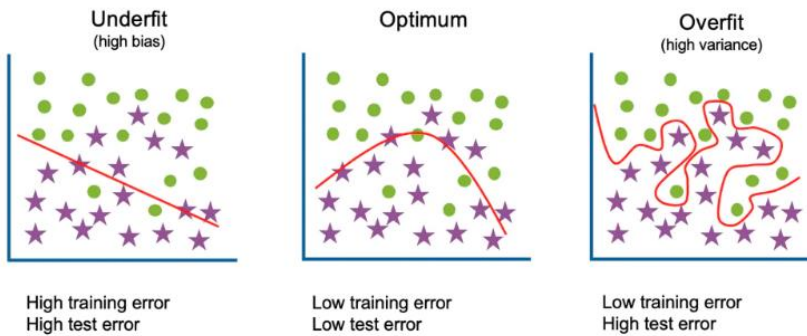
Figure 3.4: Underfitting versus overfitting [24]

To evaluate the accuracy of the forecasting model, in practice is it common to separate the available data between training and testing data [4]. The training data is used to estimate the parameters of the forecasting data, while the testing data is used to determine the accuracy. Typically, 20% of the available data is used as testing data, although it depends on the size of the data set and the time horizon of the forecast.

As the conclusion is based on a relatively small test set, it is possible that the error is not representative for future data. K-*folds cross-validation* is introduced to calculate a more robust error. For this approach, the learning data is divided over k disjoint subsets with an equal amount of data. The model is trained using k-1 subsets, which together represent the training set. Subsequently, the model is applied to the remaining subset, which is the test dataset. This procedure is repeated until each subset has served as a test dataset [26]. The forecast error is the mean forecast error of each subset. A large advantage of cross-validation is that all the available data is used to train the model. In addition, the calculated error is more trustworthy as it is based on multiple subsets, rather than just one.

As time series are the input for the forecast model, the order in which the data points are measured are very important. Suppose that the daily pick lines of April are selected as test data, then the data after April cannot be taken into account to forecast April. As no future observations can be used in constructing the forecast [4]. That is why a variant of cross-validation has been devised, which is called *time series cross-validation.* Figure 3.5 illustrates time series cross-validation for one-period-ahead forecasting. The blue dots form the training set and the orange dots are the test set. The white dots are not taken into account. The forecasting accuracy is computed by averaging the test sets [4].
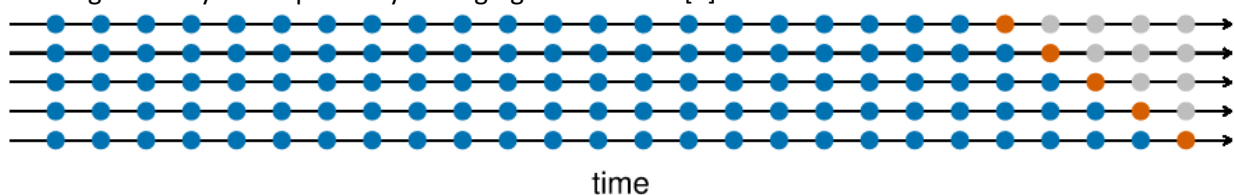
Figure 3.5: Time series Cross Validation for one-period-ahead forecasts [4]

For models that contain smoothing parameters a split in training and testing data is not sufficient to prevent overfitting. This is because the smoothing parameters have to be optimised based on a separate dataset. Hence, the time series has to be split into a training, testing and validation set of respectively 70%, 20% and 10% for the ETS and ARIMA model.

3.5 Determine the standard time per task

This section answers the sub-research question 2.4: *How can the standard time of a task be determined?*

Groover [27] defines the standard time for a given task as the amount of time that should be allowed for an average worker to process one work unit using the standard method and working at a normal pace. The standard time should include some additional time, called allowance, for the personal need, fatigue and

unavoidable delays during the shifts. Formula 3.27 depicts the formula to calculate the standard time, in which T_{std} is the standard time, T_n represents the normal time and A_{pdf} is an allowed factor for personal time, fatigue and delays. The allowance is typically set to 5% for light and medium work and 20% for heavy work.

$$T_{std} = T_n\left(1 + A_{pdf}\right) \tag{3.27}$$

Groover [27] created an overview of methods that can determine the time standards, this overview is depicted in Figure 3.6.
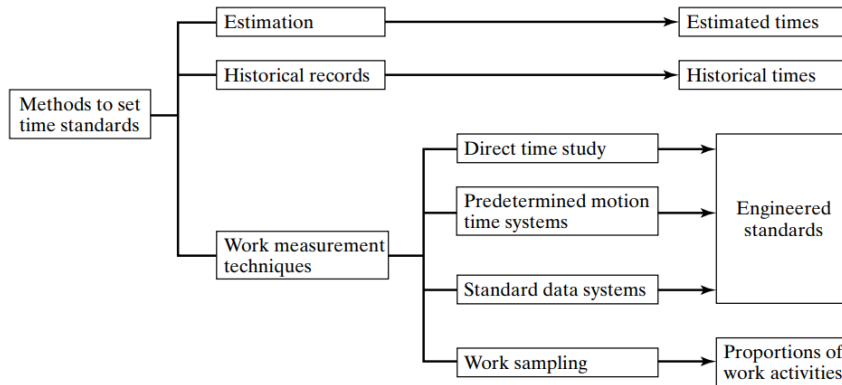
Figure 3.6: Overview of methods to set time standards

The first method is an estimation. In this case, a supervisor or foreman is asked to give an indication of how much time should be allowed for a given task. This method depends on the judgement of the estimator and is therefore the least accurate. The second method is to use historical records of identical or similar job orders to determine the standard times. Historical records are an improvement over estimation as they represent actual times for amounts of work completed. However, these records do not contain any information about the efficiency with which the work was completed [27]. The last method consists of four techniques, namely direct time study, predetermined motion time systems, standard data systems and work sampling. These four work measure techniques are used at different levels of the task hierarchy. As depicted in Figure 3.7, a job consists of tasks, which consists of work elements and finally each work element consists of basic motion elements.
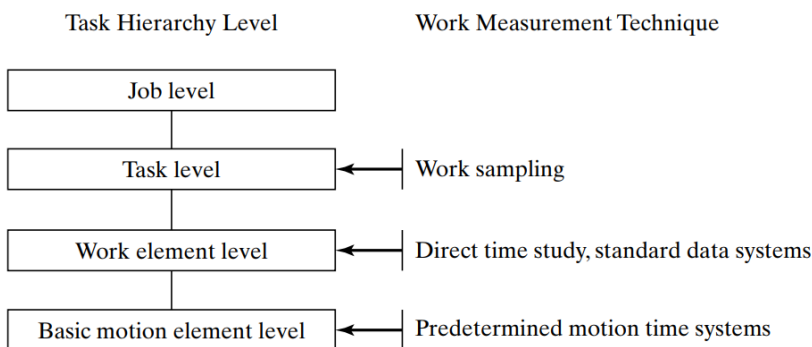
Figure 3.7: Overview of how work measurement techniques correspond to different levels in the general task hierarchy

The predetermined motion study relies on a database of basic motion elements, for example reach and move. For each movement a set of normal times is associated with it, whose values depend on the conditions under which the motion element was performed. For example, the time to move an item depends on the distance and weight of that item [27].

The data gathered in predetermined motion time systems is stored in standard data systems. The result is a database of times. The advantage is that if the standard time of a new work element has to be determined, no new study has to be carried out. All the applicable basic motion elements and their corresponding standard times can simply be added. This same method can also be used to determine the standard time of

UNIVERSITY OF TWENTE.

a task. Often the standard times of all applicable work elements are determined by means of a direct study. This is an accurate work measure technique, however a lot of time needs to be invested into building the standard data system via direct time studies and predetermined motion time studies.

A direct time study involves directly observing a worker during his task. All times are recorded via a stopwatch. Each work element is often timed separately. During the observation the worker's pace is also evaluated. The observed time is finally multiplied by the performance rating to determine the normal time for the element or the task.

The last method is work sampling, in this case a random worker and task is evaluated. The advantage for this is that, over an extended period of time, a large number of observations is made. This leads eventually to achieve the desired level of accuracy.

Due to the limited duration of this thesis, building a standard data system by means of predetermined motion time studies is not achievable. Therefore, we rely on direct time studies to determine the standard time of the tasks. The advantage of this method is that the efficiency is also observed during this study and can therefore be taken into account. The disadvantage of this method is that it is very time consuming, which means that only limited data can be collected. The WMS system stores the start and end time of trolley picking, hence historical data. For this task, it is checked whether one of the two approaches introduced by Law [28] can be used to specify a distribution which is capable of representing the standard times.

The most desirable method is to use standard techniques to fit an empirical distribution function to the historical processing times. The quality of fit is tested by hypothesis tests, for example the chi square test. In case that the theoretical distribution with certain parameters is a good model for the processing time, we then can sample from this distribution when a service time is needed in the simulation [28]. In case that no theoretical distribution passes the hypothesis test, the historical data can also directly be used in the simulation.

3.6 Scheduling

Once the most appropriate forecasting method is selected and the method is validated, then the forecasting results can be used as input for the workforce schedule [7].

3.6.1 Placing the schedule models within the control levels of an organisation

The terms planning and scheduling are often used interchangeably, however there is a difference in the time horizon. Planning is executed on the static and tactical level thus respectively on the long- and midterm. Scheduling, on the other hand, is executed on the operational level, hence on the short term. Hans, Houdenhoven and Hulshof [2] developed an overview of the planning and control framework for each hierarchical level and for each managerial area, which is depicted in Table 3.4. The most relevant managerial area for this research is the resource capacity, which is depicted in bold.

Table 3.4: Manufacturing planning and control framework based on [2]

	TECHNOLOGICAL PLANNING	RESOURCE CAPACITY	MATERIAL COORDINATION
STRATEGIC	R&D	**Demand forecasting**	Warehouse inventory management
TACTICAL	Engineering	**Resource capacity loading**	Procurement/ purchasing
OFFLINE OPERATIONAL	Micro process planning	**Scheduling**	Order picking
ONLINE OPERATIONAL	--	**Monitoring, shop floor control**	Material replenishment

Long-term forecasting is very useful for the strategic control level, planning decisions that can be made within this level are for example: redesign the process or the acquisition of expensive machines. On the tactical control level, mid-term planning can be executed. At this level, the workforce capacity is flexible. Therefore, it is possible to hire or lay off employees or change the division between permanent and flexible workers. The operational control level can be divided between the offline and online level. At the offline operational level, the schedule can be executed in advance. Within Ceva, at this level it can be decided how much workforce is required the next business day. Hence, this schedule is called the workforce schedule. However, it should be considered that the employment agency has a limited workforce pool, which can not be extended within one day. At the online operational level, decisions must be made ad hoc. It mostly involves unexpected events, like emergency orders. Therefore, the *workforce schedule* is created to support decision making at the offline operational control level, while the *online schedule* supports decision making at the online operational control level.

3.6.2 Overview of the workforce scheduling models

From the previous section, it is concluded that a scheduling tool is required that can perform well on the offline and the online operational levels. Naturally, a combination of multiple models is also allowed if they can cooperate. The focus of this section is the workforce scheduling tool. Therefore, this section answers the sub-research question 2.3: *Which scheduling models are available within the literature or used in similar environments?*

Relevant scheduling problems from the literature can be divided into three categories as depicted in Figure 3.8. The first category is *Personnel Scheduling*. If a company operates more than five days a week, the days off to which an employee is entitled must be taken into account [29]. As Ceva Heerlen operates only on business days, the *Days off scheduling* is not applicable. A second scheduling model that falls within the personnel planning category is *shift scheduling.* Companies where demand fluctuates over short time intervals face this type of scheduling problem [29]. For Ceva this is not so relevant because most cut-off times are at the end of the day. Hence, with the current morning- and afternoon shift, it should be possible to achieve all cut-off time. In addition, both variants of Personnel planning assumes that each employee needs to be scheduled, while Ceva wants to know the optimum number of employees needed. Therefore the Personnel Scheduling problems are not considered further during this research.
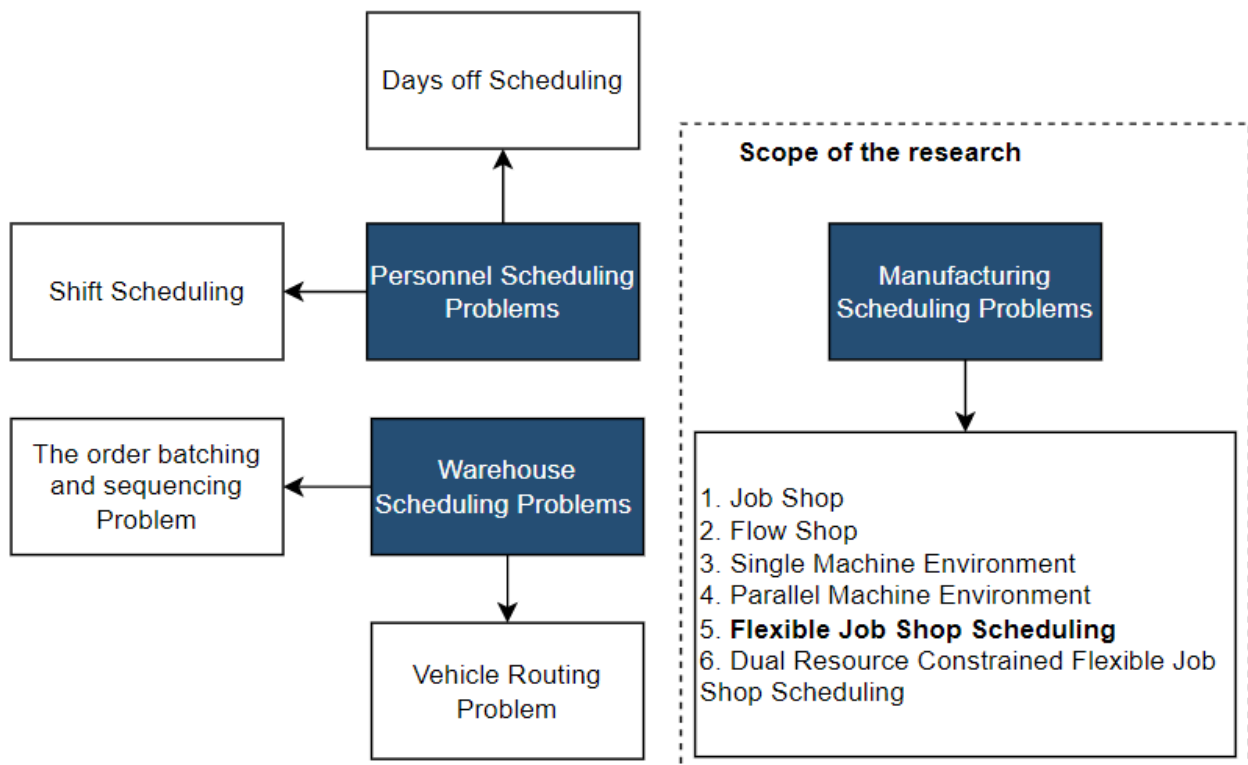

Figure 3.8: Framework of scheduling problems [29] [30] [31] [32]

The second category is the *warehouse scheduling problems*. To efficiently pick products, the travelling distance needs to be minimised. Given a set of end customer orders, fixed locations of the items to be collected and a picker router policy, the *Order Batch Problem* groups products that can be picked by a single picker within a single tour. One group is called an *order batch* [30]. As stated in section 2.6, it is assumed that the scheduling tool cannot influence the order batches which are determined by the warehouse management system. As a result, the order batch problem falls out of scope for both the workforce schedule model as well as the online schedule model.

The objective of a *Vehicle Routing Problem* is to minimise the total distance travelled by a set of vehicles serving customers spread across different locations [31]. Translating this problem to Ceva, the objective is to minimise the distance that an operator has to travel to pick all the items within a batch. However, as with the order batch problem, the order of the picks within a batch is determined by the warehouse management system and cannot be changed by the scheduling tool. Based on this, the vehicle routing problem is not considered further in this research.

The last scheduling problem that is distinguished is *Manufacturing Schedule Problems.* At first glance, manufacturing scheduling problems may not seem like the best description of a schedule problem in a logistic company. However, looking at the problem in detail, it is very similar to creating a production planning. The objective of the workforce scheduling is to determine how many employees are needed per task to meet the cut-off time of each order. As mentioned in section 2.6, order processing must be carried out according to a fixed processing sequence. Each task must be assigned to an operator who is certified for that task. In addition, tasks also require MHEs which are basically machines. Each job has its own release date, due date and processing time. Known variants of manufacturing scheduling problems are the *single* and *Parallel Machine Environment*, the *(flexible) Job-Shop* and the *Dual-Resource Flexible Job-Shop [32]*.

Single Machine Environment
This is the simplest scheduling problem within manufacturing. There exists only a single machine, which is capable of performing the operations for each job.

Parallel Machine Environment
Each job requires only one operation, which can be processed by any machine that belongs to a given subset. Machines which do not belong to the subset are not capable of performing the operation[32].

Job-Shop Schedule Problem
Each job must be processed according to a predetermined route with m jobs. The general idea behind this model is to find the optimal schedule for assigning n-jobs to m-machines [33]. Various objective functions exist for this problem, among which minimise the makespan or minimising the tardiness. The schedule problem within Ceva can be defined as a JSP. The original JSP is defined as:

Given is a set $J = \{J_1 \dots J_n\}$ of n jobs. Each Job $J_i \epsilon J$ consists of a set of $O_i = \{i_1, \dots, i_{qi}\}$ of q_i operations that have to be processed on a set $M = \{M_1, \dots . M_m\}$ of m machines [34].

However, the JSP assumes that only one machine is capable of performing a particular operation. While in practice, several identical machines are available in the warehouse. When a trolley is required for an operation, it does not matter whether trolley A or trolley B is used. The model is allowed to select any available machine within a machine type group. To include this aspect within the JSP model, the *Flexible Job-Shop Scheduling Problem* (FJSP) was introduced.

Flexible Job-Shop Scheduling Problem
The Flexible Job-Shop schedule combines the Job-Shop and the Parallel Machine Environment [32]. It can be visualised as a collection of work centres with each a number of identical machines in parallel. The job follows its own route, if an operation of the job has to be processed by one work centre, each machine within this work centre is capable of performing the operation. The scheduling problem consists of a routing sub-problem, which assigns each operation to a machine from a set of capable machines, and the

scheduling sub-problem, which sequences the assigned operations on all machines to obtain a feasible solution and minimise the makespan [35]. The FJSP is only capable of assigning operations to one resource type, so either to operators or to MHE's.

As mentioned in section 2.5.2, pick lines can be processed together in some production steps. These pick lines are either batched by trolley, box or both. A small example is that after a trolley has been completely picked, the boxes on the trolley are placed on the conveyor belt. In practice, these boxes are placed one after the other on the conveyor belt as soon as the entire trolley has been picked. The pick completion time for each box that is linked to the same trolley is therefore the same. Hence, within some production steps, the pick lines need to be scheduled in batches to represent reality. Ham And Cakici [36] introduced a Flexible Job-Shop Scheduling Problem capable of processing within batches. Their formulation is introduced in this thesis as they are capable of achieving good scheduling for large-scale problems in under five seconds [37]. A small competition time as Ceva operates in a fast changing environment, where new orders arrive on a regular basis. The sets, parameters, decision variables, resultant variables and the MIP model are described within this section.

Sets

J	Jobs (j)
S	Steps (s)
M	Machines (m)
B	Batches (b)

Parameters

$par_{j,s,m}^{ptime}$	Processing time
par_m^{capa}	Capacity of machine m
nj	Number of jobs
ns	Number of steps
L	$\sum_s MAX_{j,m}\{par_{j,s,m}^{ptime}\}$

Decision variables

X_{jsb}	1 if job j is assigned to batch b at step s, otherwise 0
Y_{bsm}	1 if batch j is assigned to machine m at step s, otherwise 0

Resultant variables

$J_{j,s}^{start}$	Start time of job j at step s
$J_{j,s}^{complete}$	Completion time of job j at step s
$M_{b,s}^{start}$	Start time of batch b at step s
$M_{b,s}^{complete}$	Completion time of batch b at step s
$M_{b,s}^{ptime}$	Processing time of batch b at step s
$Cmax$	Maximum completion time (makespan)

MIP

Routing

$$\sum_b X_{jsb} = 1 \ \forall j,s \tag{1}$$

$$nj\left(\sum_m Y_{bsm}\right) \geq \sum_j X_{jsb} \ \forall j,b \tag{2}$$

$$\sum_m Y_{bsm} \leq 1 \ \forall j,b \tag{3}$$

$$\sum_j X_{jsb} \leq par_m^{capa} + nj(1-Y_{bsm}) \ \forall b,s,m \tag{4}$$

$$X_{jsb} + Y_{bsm} \leq 1 \ \forall b,s,m : par_{j,s,m}^{ptime} = 0 \tag{5}$$

Scheduling

$$M_{b,s}^{ptime} \geq par_{j,s,m}^{ptime} - \left(2 - X_{jsb} - Y_{bsm}\right)L \ \forall j,b,s,m : s = 1 \tag{6}$$

$$M_{b,s}^{complete} \geq J_{j,s-1}^{complete} + M_{b,s}^{ptime} - \left(2 - X_{jsb} - Y_{bsm}\right)L \ \forall j,b,s,m : s = 2 \tag{7}$$

$$M_{b,s}^{complete} \leq J_{j,s}^{complete} - (1 - X_{jsb})L \; \forall j, b, s \tag{8}$$

$$M_{b,s}^{complete} \geq J_{j,s}^{complete} - (1 - X_{jsb})L \; \forall j, b, s \tag{9}$$

$$M_{b,s}^{complete} - M_{b,s}^{ptime} = M_{b,s}^{start} \; \forall b, s \tag{10}$$

$$M_{b,s}^{complete} \leq M_{b,s}^{start} + (2 - Y_{bsm} - Y_{\hat{b}sm})L \; \forall s, m: b \geq \hat{b} + 1 \tag{11}$$

$$M_{b,s}^{complete} \leq M_{\hat{b},s+1}^{start} + (2 - Y_{bsm} - Y_{\hat{b},s+1,m})L \; \forall s, m: s \leq ns - 1 \tag{12}$$

Measuring

$$Cmax \geq M_{b,s}^{complete} \; \forall j, s \tag{13}$$

The objective function of this model is to minimise the maximum completion time of the whole production schedule.

Constraint (1) makes sure that each job is assigned to a batch at each step s. Constraint (2) determines that if a job is assigned to a batch then that batch should be assigned to a machine. Simultaneously, a batch cannot be assigned more to more than one machine, which is addressed in constraint(3). Constraints (4) ensures that either the job can be added to a batch in a step or the machine can be used to process that batch in that step but not both. Constraints (5) prevents that a batch is produced on a machine which is not capable of processing all the jobs within that batch individually. Constraint (6) determines the processing time of a batch, which should be at least equal to the longest processing time of all jobs within that batch. L represents a big number. Constraints (7)-(9) determine the completion time of the batch. Constraint (10) calculates the start time based on the production time and completion time. Constraint (11) ensures that a preceding batch cannot start before its predecessor batch is completed, if they are produced by the same machine. Constraint (12) ensures that a batch can only start the next step once the preceding step is completed. Constraint (13) calculates the makespan of the schedule.

Required adjustments to implement the FJSP problem in practice
This model needs to be adjusted to implement it in practice. First of all, within this research our main focus is not to minimise the total makespan of this schedule, but to minimise the lateness of the jobs, which is also often called the tardiness of the jobs. To monitor it in the first place, we need to add a parameter that represents the deadline of each job. In literature, this is often represented by DD_j, which means the due date of job j. The tardiness per job is the difference between the completion time and the due time, assuming that the due date and completion date are the same.

Another difference between the model described by Ham and Cakici [36] and the reality is that in the model it is an option whether the product is produced within a certain batch or individually. While at Ceva Heerlen this is not optional, a trolley picker is not allowed to carry a single box all the way to the conveyor belt every time all the products for that box are picked. Therefore, the model has to be adjusted so that if a pick line is part of a batch, it must be produced as a batch.

After these adjustments the FJSP model is capable of minimising the tardiness of the jobs by assigning the operators optimally to the different tasks, while taking into consideration the batches and the order in which the pick lines have to be processed. The model assumes that there is a limited number of operators available. This makes the model extremely suitable for the future online scheduling application, which fits well with Ceva's strategic goals. However, for the offline workforce scheduling, the goal is to determine the number of operators needed per task. Thus there is no limited number of operators in this phase of the research yet.

Therefore, the FJSP model in the workforce scheduling tool is used to calculate how many pick lines can be completed on time and how many pick lines are processed after the deadline. The FJSP model uses as one of its input parameters the number of operators for each task that is available to process the pick lines. The

great advantage of the implementation of the FJSP model is that reality is simulated well and the actual processing order is respected. By adjusting the input parameters, and thereby thus the number of operators scheduled per task, the influence at the tardiness and idle time can be measured. The FJSP model needs to be calculated several times for the workforce scheduling model, with varying input parameters in order to find the best workforce schedule.

Dual-Resource Constraint Flexible Job-Shop problem

A *Dual-Resource Constraint Flexible Job Shop Problem* (DRCFJSP) is capable of simultaneously dealing with a flexible selection of machine and worker. Andrade-Pineda, Canca, Gonzalez and Calle [37] included the makespan and due-date related criteria in the DRCFJSP model they formulated. This model ensures that a machine and a worker are always simultaneously available to perform the production step. However, in this case it is not necessary to apply the DRCFJSP model, as operators for a certain task always work with the same MHEs. For example, only a reach operator uses a Reach Truck. And even more importantly, he uses this Reach truck permanently during his shift. The number of available MHEs is therefore equal to the maximum schedulable number of operators for that task. For example, it is not necessary to schedule more Reach operators than that there are Reach trucks available. This means that the number of Reach trucks available is equal to the number of Reach operators that can be scheduled at the same time. Hence, a maximum number of schedulable operators per tasks must be defined as an additional constraint for the Flexible Job-Shop Problem as described by Ham and Cakici [36].

3.6.3 Complexity of the problem

The Flexible Job-Shop Problem is an extension of the JSP. To determine the optimal solution of a JSP, all $(n!)m$ possible schedules need to be calculated. The computation time will therefore increase fast as the number of jobs increases. The JSP is one of the most well-known models that is NP-hard and thus computationally intractable [38]. Jain and Meeran [38] created an overview of methods that are able to solve the JSP. The most common methods are depicted in Figure 3.9.
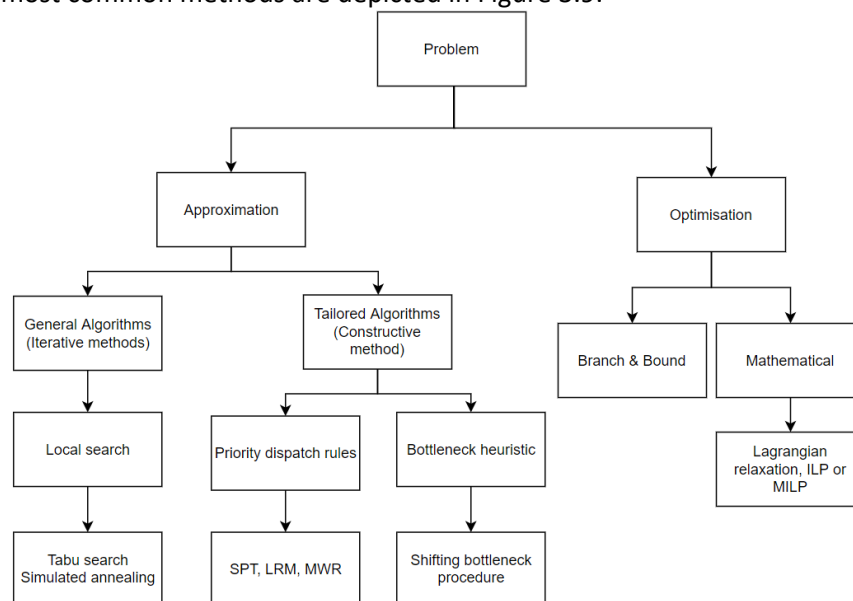
Figure 3.9: Selection of overviews to determine a (near) optimal job shop schedule [29]

A problem can either be solved exactly or approximately. Exact algorithms that optimally solve the problem are, among others, branch and bound, integer programming and lagrangian relaxation. Cheng and Matis state that using an exact method to solve the JJSP is not practical for large-scale instances [39]. Considering the large number of jobs that need to be scheduled to determine the number of operators needed for Ceva, the use of a heuristic is a good alternative. A heuristic, also known as an approximation model, is capable of finding a near-optimal solution in less computation time. As is shown in Figure 3.9, the heuristics that are capable of solving the JJSP can be divided in iterative and constructive heuristics.

Constructive heuristics start with an empty solution space and add one job to the schedule within each iteration. Improving heuristics starts with a complete schedule and tries to improve the schedule further by changing the order of jobs. Improving heuristics, like tabu search and simulated annealing are likely to suffer from scalability issues for large problems. This is especially not practical for rolling schedules [39]. As a result, the constructive heuristics *priority dispatching rule* and *the bottleneck heuristic are explained* in more detail.

Priority dispatching rules assign a priority to jobs that are waiting in the queue, based on a combination of the jobs characteristics and common system level parameters, the job with the highest priority is then selected [39]. Some well-known priority rules are: First in first out (FIFO), Shortest Processing Time First (SPT), Longest Processing Time first (LPT) and the Weighted Shortest Processing Time first.

In addition, it is also allowed to combine priority rules as they often perform better than a single rule. The shifting bottleneck is a powerful heuristic to solve the JSP [40]. In each iteration the model searches the bottleneck machine, which is the machine with the longest makespan. This machine is added to the schedule. Subsequently all the machines that are already in the schedule, are recalculated in order to reduce the makespan of the schedule. This procedure is continued until all machines are scheduled.

3.6.4 Online schedule

This section answers sub-research question 2.5: *How can online scheduling be implemented within this model to adjust the schedule based on real-time data?*

Scheduling purely based on the predicted productivity is unrealistic and results in deviations of the production schedule [41]. To create a more accurate schedule Heger, Grundstein and Freitag [41] have introduced a method to include past and real-time data within online scheduling. This approach is depicted in Figure 3.10.
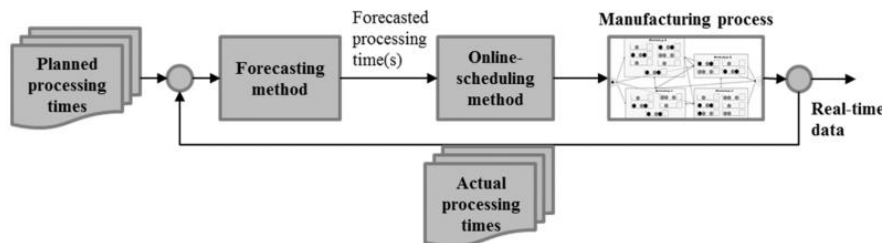
Fig. 1. Basic approach.

Figure 3.10: Online-scheduling using past and real-time data [41]

The model starts by scheduling based on the planned processing times. Once the employees start their shift, the planned productivity can be adjusted based on real-time data. The deviation between the planned and actual productivity is presented by productivity deviation factor $d_{i,k,l}^t$. Every time a job is finished at any station, the $d_{i,k,l}^t$ is determined via equation 3.27. The value is calculated for each operator i, shift k, station l. $pt_{o,i,k,l}^t$ represents the productivity of period t, o represents the operation and $pt_{o,i,k,l}^{plan}$ represents the planned productivity.

$$d_{i,k,l}^t = \frac{pt_{o,i,k,l}^t}{pt_{o,i,k,l}^{plan}} \tag{3.27}$$

The $d_{i,k,l}^t$ is used to forecast the productivity factor for the next job, which will subsequently be used to adjust the schedule. The method is a general approach, hence it is not restricted to a certain forecasting or scheduling method. An example is given by using the weighted average form of simple exponential smoothing. As shown in formula 3.28, the forecasted productivity factor $\hat{d}_{i,k,l}^t$ is equal to a weighted average between the most recent observation $\hat{d}_{i,k,l}^t$ and the most recent forecast $\hat{d}_{i,k,l}^{t-1}$. Next, the forecasted productivity can be calculated using formula 3.29.

$$\hat{d}_{i,k,l}^{t} = \propto * d_{i,k,l}^{t} + (1-\propto) * \hat{d}_{i,k,l}^{t-1} \tag{3.28}$$

$$\widehat{pt}_{o,i,k,l}t = pt_{o,i,k,l}^{plan} * \hat{d}_{i,k,l} \tag{3.29}$$

Heger, Grundstein and Freitag [35] advise to use a simple forecasting method as the ES forecasting model did not outperform the Naïve forecasting. Using more complex forecasts will lead to longer computation times, which is challenging to implement in a real-time model.

Ultimately, a more accurate schedule can be based on the adjusted productivity rate and by including the newly arrived pick lines. As the day progresses, the cut-off times get closer, the schedule can advise the supervisor to reallocate operators to bottleneck departments so that the backorders are minimised. The optimal number of operators per department, considering that the total number of operators is fixed, could possibly be determined via the DRCFJSP as well. The computation time of the model is very important here, as this model needs to be calculated multiple times during the day. In addition, it will only be of added value to the supervisor if the schedule is based on real-time data. Data from a few hours ago is not so relevant, as situations can change very quickly within this sector. If the computation time of the DRCFJSP proves to be too long, a simplified version should be used.

3.7 Summary of the literature review

This chapter introduced multiple forecasting methods that can be used to forecast the next day's pick lines per MHE. As mentioned the ETS function, Arima function, regression based analysis model and the artificial neural network models are evaluated further in Chapter 4. In addition the Naïve forecast, which assumes that the predicted value equals the last measured value, is used as a benchmark. By comparing the forecast error of the evaluated models with the Naïve forecast error, we can determine the added value of the forecast model in terms of improved accuracy.

Additionally, the Flexible Job-Shop Schedule Problem was introduced which is capable of determining the maximum makespan of the jobs and the mean tardiness. By running the model multiple times with a varying number of operators, the optimal number of operators can be determined.

Subsequently, by using online scheduling, the productivity can be adjusted throughout the day in order to match reality as precisely as possible. By recalculating the FJSP model regularly, the newly arrived orders are added to the schedule. This will result in an up-to-date schedule.

4 Model description

The previous chapters described the problem situation and introduced useful models from literature. In this chapter, this knowledge is combined into a forecast model and a workforce scheduling model that answer the main research question:

How can workforce scheduling of operational employees be improved at Ceva Logistics Heerlen by including short demand forecasting and estimation of standard times?

The forecast model solves one of the identified core problems, namely that currently no accurate forecast model exists to predict the next day's pick lines per MHE type. The literature review showed that the ETS, ARIMA model and linear regression based models are the most promising for this application. Section 4.1 describes in detail the method to determine the most appropriate forecast model per MHE type.

Subsequently, the standard time per task is determined in section 4.3.1 via direct time studies and historical data. The standard times per task and the forecasted number of pick lines are the input for the workforce scheduling model which is introduced in 4.3. The workforce schedule model solves the '*Conversion from the demand to required hours is missing*' core problem.

4.1 Forecasting Model

This section describes the implementation of each forecast model, which is visualised in a flowchart depicted in Figure 4.1. According to the forecast methodology [7], as described in chapter 3, the first step is the historical statistical data collection. The raw data is collected from the WMS system, this is described in Section 4.1.1. Section 4.1.2 describes in detail how the extracted data needs to be analysed and verified before it can be used as input. This involves filtering the data, removing weekends and replacing the public holiday data. Subsequently, section 4.1.3 shows the results of the preliminary analysis that is performed on the historical daily pick lines. Based on this, it was decided to implement the ARIMA model twice, with deviating time series seasonal frequencies. To distinguish both ARIMA models in the flowchart, the models are called ARIMA1 and ARIMA2 for convenience.

In order to determine the optimal models and to validate the results, the data is split into a training set and a test set for the Naïve and linear regression model. For the ARIMA models and the ETS model, a three-split is required in a training, test and validation set. In section 4.1.4 it is explained why splitting the data by two different methods is necessary to obtain the optimal results.

Section 4.1.5 introduces the implementation of the simplest forecast model, the Naïve model. The implementation of the ARIMA models and the ETS model are described in section 4.1.6. As shown in the flowchart, the training set is used to determine the model parameters while the test set is used to optimise the smoothing parameters.

All the steps mentioned until now are depicted in black in Figure 4.1. All models discussed so far are individual models. The purpose is to compare whether the ETS, ARIMA or the Naïve model is capable of predicting the pick lines the most accurately for each MHE type. However, as mentioned in the literature review, Hyndman [4] states that an easy way to improve the forecast accuracy is by using several different methods on the same time series and averaging the results. Hence, a combination of time series forecast models often leads to a smaller forecast error than the forecast error achieved by the component model that performs best. Based on this knowledge, we have chosen to create a combination forecast between the Naïve model, the ETS model and both ARIMA models. The additional step that is used exclusively for the combi model is depicted in blue within the flowchart. This combi model uses the results of the Naïve training set, the ARIMA and ETS test sets and determines an optimal weight for each of the individual components to calculate the weighted average forecast. This method is further elaborated in section 4.1.7.

As briefly introduced in section 2.4, the aggregated pick lines can also be used to forecast the expected aggregated pick lines of all MHE types. Just like for the individual MHE type time series, also for the

aggregated time series the optimal model and smoothing parameters must be determined for the ARIMA models and the ETS model. Next, the aggregated time series must be split over the MHE types by determining the weight with which each MHE type contributes to the aggregated daily pick lines. This method, also called the Top-Down Approach is further explained in section 4.1.8. In the flowchart, the extra steps that are only required for the Top-Down Approach are shown in pink. It is possible to use the Top-Down approach for the Combi forecast model.

Finally, section 4.1.9 describes the linear regression model. For this model not only the daily pick lines are required as input but also the known pick lines. The aim of the linear regression model is to determine a formula that describes the relationship between the two time series. The linear regression formula is determined based on the training set by minimising the least squared error.
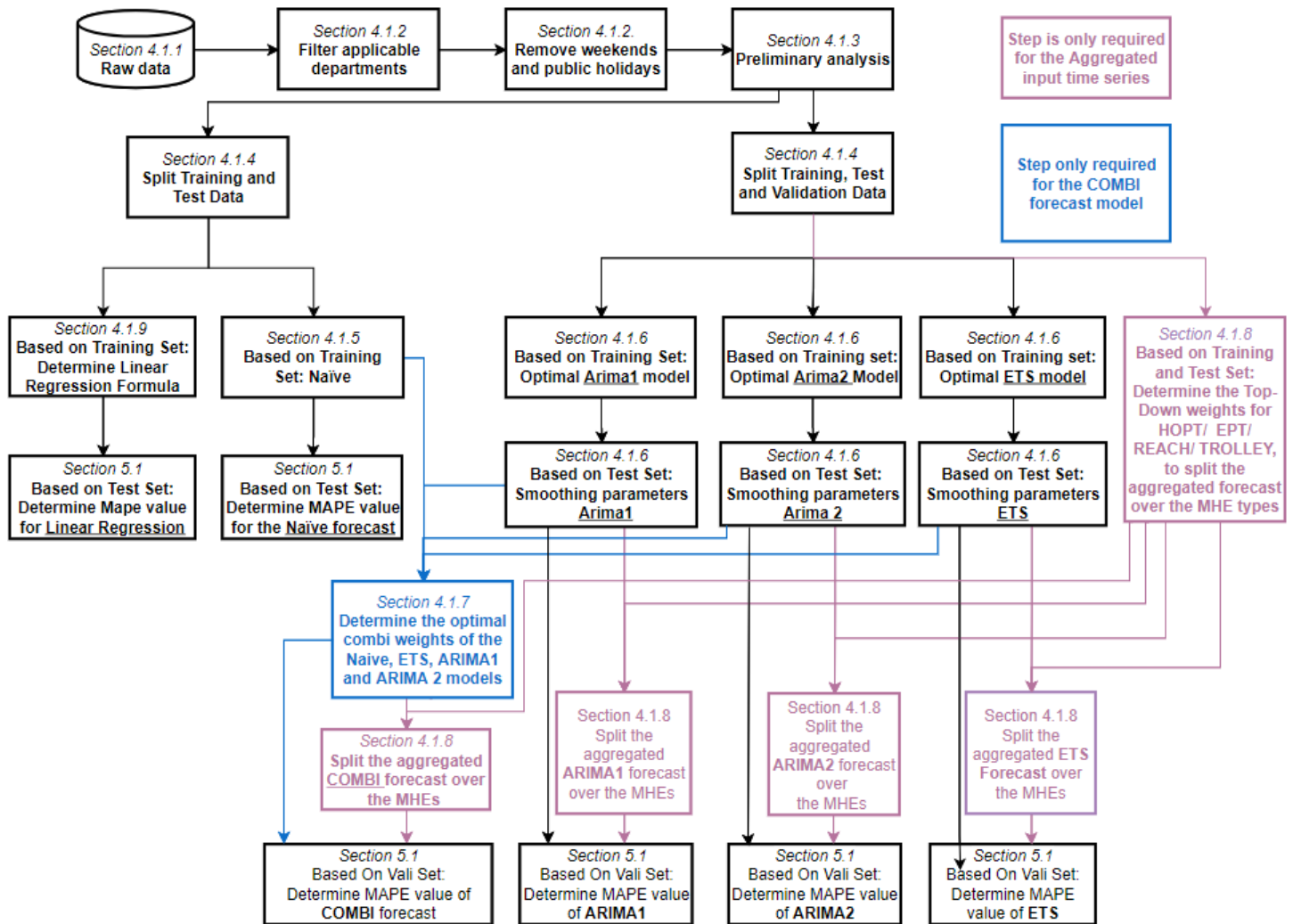
Figure 4.1: Flowchart to determine the most suitable forecast model per MHE type

Each model is validated with a separated data set, which is unknown to the model to prevent overfitting. For the linear regression model and the Naïve model the test data set is used, while for the ARIMA and ETS model the validation data set is used. The performance of each model is expressed in the MAPE value. This is discussed in Chapter 5 where the results of each model are also presented.

4.1.1 Data collection

All activities within the warehouse are saved in the SAP system. The historical data remains available for two years before being deleted automatically. Performing data extractions from SAP, results in multiple separate tables. For example, every order that comes in, is saved within one table of SAP together with the timestamp of the order. Subsequently, the order is divided into pick lines which contain, among others, the

information of the responsible departments and the required skills. This information is saved in different tables. These separate tables have to be merged and linked with the right relationships-entities, to extract the applicable information. During this research, Power Query within Microsoft Excel was used for this. This choice is made because Power Query is able to perform the same data transformation for new data stored in the same folder at the push of a button. In addition, Power Query is regularly used within the Ceva, which makes it easier to implement as some staff members already possess knowledge about the tool.

After the raw data has been linked to each other, multiple filters must be applied to keep only the relevant data. For example, filtering must be done to only obtain data of the considered department. By aggregating the filtered data on a daily level, the first type of quantitative data can be extracted, which are the *Daily Pick Lines* of the past two years. The daily pick lines are known per MHE and for the whole department. Additionally, the *Known pick lines* before the scheduling moment, can be calculated daily for each MHE. Hence, the data retrieving frequency is once each business day. Both values need to be added to the corresponding time series, before the forecast model can be executed.

4.1.2 Data analysis and verification

This section answers the sub research question: *How should the data be prepared before it can be used as input data for the model?* Figure 4.2 depicts the daily pick lines at the aggregated level. One can observe directly that the pick lines show varying demand. Days on which the number of pick lines are near or equal to zero, are mainly weekends and public holidays. The daily pick line series of the individual MHEs show similar behaviour.
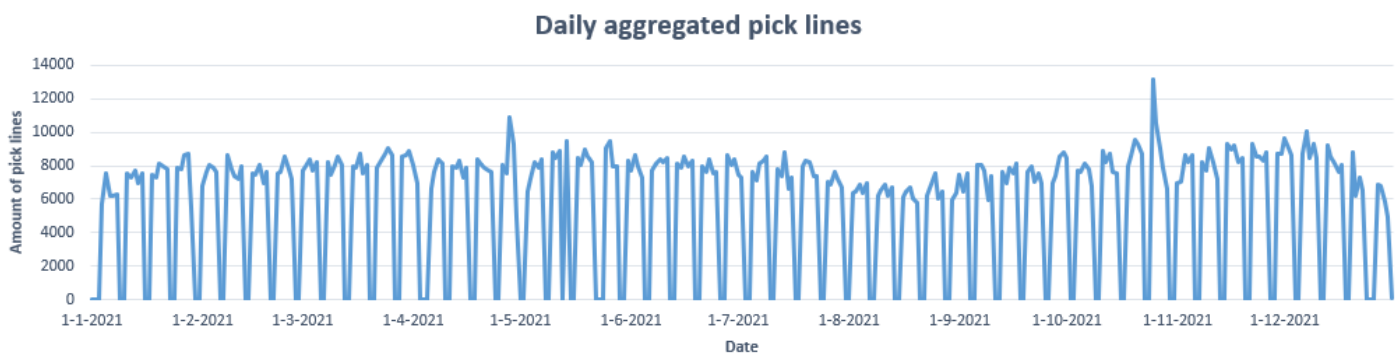

Figure 4.2: Daily aggregated total pick lines, within the year 2021

End-customers cannot select a Saturday or Sunday as shipment day. As a result, the number of pick lines that have to be processed during weekends is equal to zero. This can be included in the forecast in two ways. Either a model is used that can distinguish daily seasonality and sets the prediction value on weekends equal to zero. Or the weekends can be removed from the time series. It was decided to remove the weekend from the data in order to keep the forecast model as simple as possible as more complex models tend to have longer computation times.

Table 4.1: Total Pick Lines of Remembrance and Liberation Day for two consecutive years

PUBLIC HOLIDAY	2020	2021
REMEMBRANCE DAY	926	7913
LIBERATION DAY	1210	8948

Dealing with public holidays is more complex. In general, public holidays are particularly hard to forecast as the historical data can only be collected once a year [42]. Hence, a lot of historical data is needed to accurately forecast these days. As there is only two years of data available, this is a challenge. In addition, the client has a major influence on the number of pick lines as they determine per public holiday whether the order system will remain open or will be closed. Table 4.1 gives an example of this situation, for the Remembrance Day and Liberation Day of the consecutive years 2020 and 2021. In this case, using the pick lines of 2020 to predict the pick lines of 2021 would lead to very poor results. Therefore, we decided to

exclude public holidays from the forecast model. Huber and Stuckenschmidt [42] state that demand deviates not only on the public holiday itself, but also the day before and the day after the holiday. This is also apparent from the historical data as a peak in orders follows after a 'closed' public holiday.

To forecast the amount of next day's pick lines, the data is converted into a time series in the program R. This is necessary to be able to use the many analysing functions that are available in R. Within the program this is done by entering the start date and the frequency as shown in Formula 4.1. R will then automatically link the data to the selected time intervals. For example, if daily data contains yearly seasonality the time series frequency is 365 days.

$$Total <- ts\Big(Sales[,2], frequency = 365, start = decimal\_date\big(ymd(2020\text{-}04\text{-}05)\big)\Big) \qquad 4.1$$

Removing public holidays and weekends from the data, will also influence the time series frequency. If the weekends are excluded from the data, each week contains only five data points. Hence, each year contains 5*52=260 data points. These data points can still be transformed to a time series as each week still contains a constant number of data points.

However, when the public holidays are extracted from the data, this will result in weeks with a varying number of data points. As a result, it is no longer possible to determine weekly seasonality. An example of this is depicted in Figure 4.3. On the left hand site a section of the time series is depicted, the time series automatically assigns the day to the data. Thursday, which is depicted in Blue, is a public holiday. When this datapoint is removed, the time series will assign the next data point to 'Thursday', while in reality it is a Friday. All the data points from this moment on are shifted, hence the weekly seasonality is shifted as well from this point onwards.

DATA	WEEKDAY			DATA	WEEKDAY
8510	Monday			8510	Monday
7122	Tuesday			7122	Tuesday
8887	Wednesday			8887	Wednesday
0	Thursday			8798	Thursday
8798	Friday			8893	Friday
8893	Monday				

Figure 4.3: Result of removing values with a time series

To avoid this problem, the public holiday values can be changed to *NA* instead of removing it, meaning that the value is not available. As this value is no longer a number, it cannot be taken into account for the forecast model. For some forecast function in R this solution works as these are able to deal with *NA* values, e.g. the Arima function. However, functions like Naïve and ETS do not know how to handle these values and will therefore return an error message.

For these functions, the public holiday is substituted by the average value of that particular day. This method was chosen because it affects the weekly seasonality the least. The same method is used for the preceding and succeeding day.

4.1.3 Preliminary analysis for trend and seasonality using time series

The statistical forecast starts with a preliminary analysis, in which the time series are plotted and trend, seasonality and noise are identified. The preliminary analysis answers the sub-research question: *What causes of deviations in demand can be distinguished from the available data?* In case seasonality affects the time series, it identifies additionally whether the seasonal effect is weekly, quarterly, annually or a combination. The preliminary analysis is done for EPT, HOPT, Reach and Trolley pick lines. As mentioned in section 2.3 Trolley pick lines account for 94% of the total pick lines, therefore we have chosen to describe the preliminary analysis of the trolley pick lines in detail in Appendix B.

The conclusion of the visual preliminary analysis is that seasonality can be distinguished for the historical Daily Pick Lines for the MHE type Trolley. Of all the models evaluated during this research, the ETS model and the ARIMA model are capable of finding seasonality in a dataset. Since weekly, quarterly and annual seasonality emerges from the preliminary analysis, it is worth noting that the ETS and ARIMA functions in R have a limit of how many seasonal periods they can handle. The ETS function is able to handle a maximum seasonal frequency of 24. Meaning that ETS is only able to recognize weekly seasonality, as a quartal has already a seasonal frequency of 65 (13 times 5). The ARIMA function is, in theory, able to handle 350 seasonal periods. However, it usually runs out of memory if the number of seasonal periods exceeds 200 [43]. Hence, the ARIMA model is not able to include yearly seasonality in the model. As weekly and quarterly seasonality are possible, we have chosen to analyse two ARIMA models. The *ARIMA single seasonal period* which can only distinguish weekly seasonality and *ARIMA multi seasonal period* which can distinguish both weekly and quarterly seasonality. Both models are based on the same time series data, but have different seasonal frequencies.

4.1.4 Split in train, test and validation sets

After the data is prepared as described in section 4.1.2 the data is divided into a different set to train and test the models. For the linear regression and Naïve model a two split in only training and test data is sufficient as both models do not contain smoothing parameters. Since the ARIMA and ETS model do consist of smoothing parameters, which increases the risk of overfitting, a three split in training data, testing data and validation data of respectively 70%, 20% and 10% is required.

We chose to use the same data set to validate the linear regression and Naïve model as is used for the statistical forecast models. Hence, the last 10% of the available data is used to determine the MAPE value, to make the comparison between the different forecast models as accurate as possible.

4.1.5 Naïve model

The Naïve forecast model is the simplest model to implement. It assumes that the predicted value equals the last measured value. Hence the number of tomorrow's pick lines will equal today's pick line. As we removed the weekend in the time series, the Monday volume will be expected to equal the volume of last Friday.

There are two reasons to implement this model. As explained briefly in section 4.1, the Naïve model is one of the components for the combi model. The model has to be implemented to assess the accuracy of the model and later on determine the compound weight.

In addition, this model is used to put the accuracies of the other models into perspective. All other models are more difficult to implement and therefore the implementation is more time consuming. To determine whether this invested time actually leads to a result that is of added value. It is common to compare the forecast errors of the various methods with the error of the Naïve model.

4.1.6 ARIMA and ETS model

As shown in Figure 4.1, the first step to forecast the individual MHE types is to determine the optimal ETS and ARIMA parameters based on the training set. This can be done via the built in functions ETS and AUTO.ARIMA in R, which minimises the AIC value introduced in section 3.3.1. Globally, the ETS function calculates the AIC value for any combination of the trend and seasonality parameters. The pseudo code of the ETS function is shown in Algorithm 1 in Appendix C. For the AUTO.ARIMA function it is more complicated as there are six parameters that need to be optimised. As a result, the number of possible combinations increases rapidly. Often it is not feasible to fit every combination to simply choose the model with the lowest AIC value [44]. Therefore, a step-wise procedure for traversing the model space, Algorithm 2 in Appendix C shows the Pseudo Code of the procedure. The result of this step is the optimal ETS model and ARIMA models for the daily pick lines time series for each MHE type and ARIMA type. For the remainder of the forecast model selection, these model parameters are assumed fixed.

Each model has its own smoothing parameters based on the selected model. These smoothing parameters are parameters whose value is used to control the learning process. For example, within exponential smoothing the α value determines what weight is given to the most recent observation. The optimal smoothing parameter(s) are determined based on the test set. Similar to finding the right model, the ETS function and the ARIMA function searches for the smoothing parameter value that leads to the lowest AIC. For the individual MHE level time series, the model accuracy can be determined directly over the validation set and be expressed as the MAPE value. The model parameters and smoothing parameters are assumed fixed for the validation set to prevent overfitting.

4.1.7 Combination forecast

Hyndman [4] states that combining multiple forecasting methods tends to result in smaller forecasts errors than the individual models. Therefore an additional *combination model* is created in R, which answers the sub-research question: *How can the forecast models described in literature be modified to support Ceva in the best possible way?*

For the combination model, R is allowed to combine the four models as shown in formula 4.2. The letters a till d represent the weighting parameters that determine the weight of each forecast model. The sum of a till d should be equal to one.

$$Forecast\ value = a * \text{Naïve} + b * ETS + c * Arima + d * Arima\ multi \qquad 4.2$$

R has a function called mixture which is able to determine the optimal weight of the individual models to minimise the error of the combi model. The function starts by assigning an equal weight to each forecast model. Since the sum of the weights has to be equal to one, each individual weight starts at 1/4. During the test set it learns which model results in the lowest errors and adjusts the weights based on that knowledge. At the end of the test set, the mixture function has determined the 'optimal' weights that lead to good results for the entire test set. These optimal weights are assumed fixed at the end of the test set.

As with overfitting the smoothing parameters, it must be prevented that the 'optimal' individual weights only lead to a good result for the test data, but lead to poor results for unseen data. Therefore, the validation set is used to validate if the combi model leads to small forecast errors for data that was not seen before.

4.1.8 Top-down approach

As discussed earlier, the Top-down approach is a variant of the statistical forecast of the individual level. The difference is that not the time series of the individual time series is used as input data but the aggregated time series. The parameters of the ETS and ARIMA model are selected to minimise the AIC of the aggregated level. Once these parameters and smoothing parameters are determined an additional step is required. The average contribution of pick lines per MHE to the total pick lines must be determined. This is done based on the training and test set, which contains a period of 1.5 year. EPT, HOPT, REACH and Trolley contribute respectively 2.6%, 0.8%, 1.8% and 94.8% to the total pick lines.

As the forecast models will only forecast the total pick lines in the Top-down approach, the total number of pick lines divided over the MHE types based on their contribution percentage. Hence, the forecasted trolley pick lines are 94.8% of the forecasted aggregated pick lines.

4.1.9 Linear regression

Through the Linear Regression Based model it is tested whether there exists a relationship between the known pick lines and the daily pick lines per MHE. As by the time the forecast model has to run, at 15PM, thousands of pick lines are already known, the possibility of a relationship between the two variables cannot be ruled out in advance. Since the expected number of daily pick lines must be predicted with the

number of known lines, the number of known lines represents the independent variable and the number of pick lines represents the dependent variable.

Figure 4.1 shows the implementation of the linear regression based model. The *lm* function in R looks for a line that fits the relationship between known pick lines and daily pick lines within the training data in the best possible way. To check whether the formula found by the *lm* function also yields accurate results for new data, a split between training and test data is also used. Because this model only looks at simple linear regression, and therefore no hyperparameters are involved, a split between training and test data is sufficient. We chose to use the same data set to validate the linear regression as is used for the statistical forecast models. Hence, the last 10% of the available data is used to determine the MAPE value, to make the comparison between the different forecast models as accurate as possible.

4.2 Using the forecast as input for the workforce schedule

This section answers the sub-research question: *How can the output of the forecast model be used as input for the workforce scheduling model?*

The forecast model as described in section 4.1, will select one out of the ten introduced forecast models that is able to perform the point forecast the most accurately. The result is four forecast models, each capable of point forecasting one of the four MHE types. This point forecast is used as input for the workforce scheduling model. The goal of the workforce scheduling tool is to minimise the number of operators per task while still being able to process all the pick lines on time. At 3PM, the workforce schedule is generated for the next business day. Figure 4.4 depicts a timeline to show how the point forecasts are embedded in the workforce schedule. The timeline runs from 1 PM, the moment the evening shift starts, to the end of the last shift on the next business day.

As mentioned in section 2.2, Ceva operators work in a morning and an evening shift, both consisting of eight working hours. When the schedule is made at 3PM, the shift in operation is processing the pick lines with today's PGI date. The time at which the evening shift has processed all pick lines with today's PGI date can be determined relatively accurately since the workforce capacity and between 95-98% of the demand is known. Once all pick lines with today's PGI date have been processed, the same shift will start to process known pick lines with next day's PGI date. The time span that is still available in that shift and the number of known pick lines as demand are used to determine the number of known pick lines with next day's PGI date cannot be processed today. These unprocessed known pick lines are the known demand for the next day.
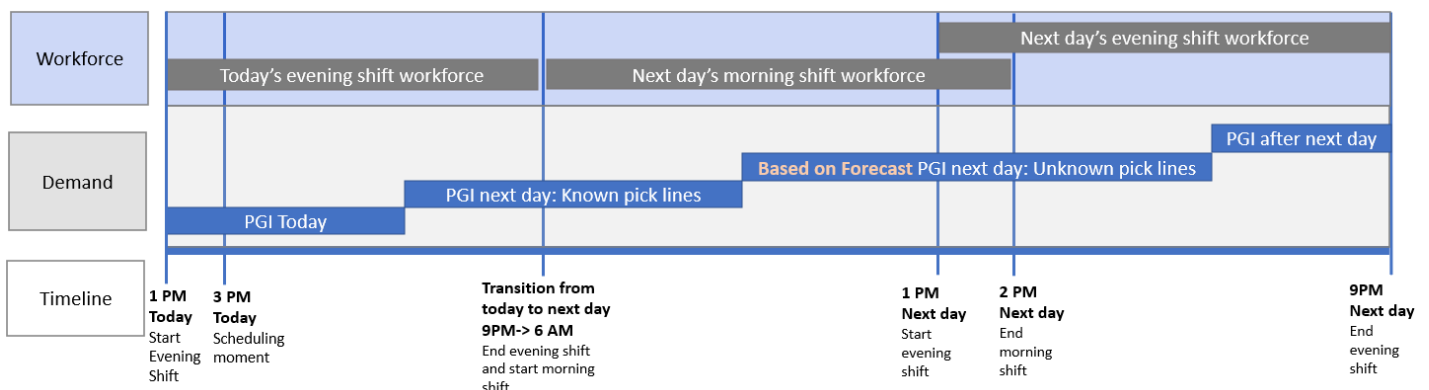

Figure 4.4: Timeline to show how the point forecast is embedded in the workforce schedule

The time at which the known pick lines with next day's PGI date are completely processed depends on the number of FTE that is scheduled per task. On a normal day, the morning shift will complete all known pick lines before the end of their shift as depicted in Figure 4.4. Each shift works nine hours a day, this includes a one hour break. Therefore, in Figure 4.4 shifts are indicated as eight hours, as this is the productive time window. The morning shift works from 6AM to 2PM and the evening shift works from 1PM to 9PM. Hence

there is one hour overlap between both shifts in the afternoon. The time span between processing the last known pick line and the end of the morning shift is used to process the unknown pick lines. These are the pick lines that are unknown at the moment the workforce schedule is made but as the next day progresses, these pick lines arrive and can be processed.

The expected number of unknown pick lines are calculated by formula 4.3 and is based on the point forecast.

$$Unknown\ pick\ lines = Point\ forecast - Known\ pick\ lines \qquad (4.3)$$

The makespan to process all the unknown pick lines depends on the number of scheduled FTE per task but also on the accuracy of the point forecast. Meaning that if just enough FTE are scheduled to process all pick lines with next day's PGI date by 9PM the next day, a higher order arrival will quickly result in backorders. Therefore the target is that at 6PM each day, all pick lines with a PGI date of that day have to be processed.

4.3 Workforce scheduling model

This section answers the sub-research question: *How should the workforce schedule model be implemented to fit the requirements of Ceva in the best possible way?*

The literature review in section 3.6 has shown that the Flexible Job-Shop Scheduling Problem is best capable of simulating the processing of orders closest to reality. This is because the FJSP model takes into account the order of the process steps in which pick lines must be processed. As the FJSP is a NP-hard problem, it is not possible to solve this problem optimally in polynomial time. Therefore a combination of two heuristics are used to find a near optimal solution. An overview of how these heuristics cooperate to solve the workforce schedule as a FJSP problem is given in Figure 4.5. Combining those two heuristics represent reality as best as possible. The reason that we do want to solve the FJSP problem as close as possible to reality instead of optimally is that it is used to determine the optimal workforce schedule. The goal of Ceva is to have a realistic estimate of how many lines can be processed with the scheduled workforce capacity.
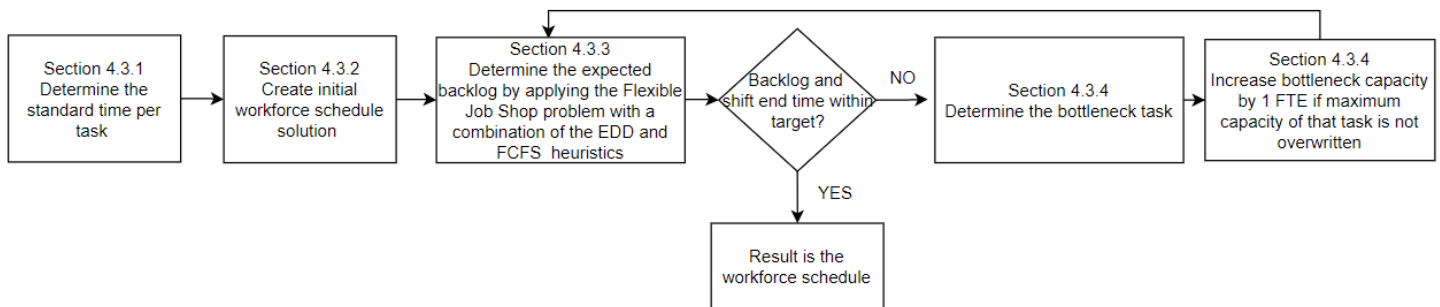

Figure 4.5: Flowchart Workforce Scheduling Tool solved as Flexible Job Shop Problem with two heuristics

The first step is to determine the standard times of each volume-dependent task. Section 4.3.1 describes in detail how the standard times are determined. Subsequently, Section 4.3.2 described how an initial workforce solution is determined. This is necessary as input for the FJSP problem, as a limited number of operators is assumed. The initial workforce solution will contain the minimum number of FTE required per task to process all pick lines in time. Section 4.3.4 explains how the expected backorders and expected finish time of each operator is determined based on the initial workforce schedule, which are the objectives of the workforce scheduling tool. To simulate reality as accurately as possible, two priority dispatch rules are used, named the *Earliest Due Date* (EDD) and *First Come First Served* (FCFS) rule. As long as the expected backorders and finish time are not within target, the workforce schedule needs to be increased by one FTE in each step. Section 4.3.5 explains how the task is determined that determines the speed of the entire process, also referred to as the bottleneck task. In addition this section describes how the initial

workforce schedule is updated. Then the expected backorders and finish time is determined again for the updated workforce schedule via the FJSP. This procedure is repeated until the target is reached.

4.3.1 Standard times of the volume-dependent tasks

An important step to create a valid workforce schedule is to determine standard time per task. As described in section 2.2, there is a lack of insight into the standard times. Only for the trolley pickers it is entered how many pick lines they have to process per hour. However, this is not a tangible value for the picker as he does not know the number of lines linked to the box or trolley that he processes. Furthermore, not only the trolley picker task has a standard time but also the other volume-dependent tasks. Therefore this section will answer the sub-research question: *What variables influence the total standard time of a pick line?*

The standard times of eight of the twelve tasks have been determined through direct time studies. This means that we directly observed and timed these tasks. This choice was made because no data was available yet. Direct time studies are a relatively fast and reliable way to determine standard times. By this method, not much data is collected, but external circumstances, such as disturbances, can be included as they are also observed and noted. For all tasks a driver is identified upon which the standard time is based. The overview of the standard times is given in Table 4.2. Due to compelling business interest, it is not possible to state the actual standard times per task in this thesis. Ceva Heerlen can extract this information from the model parameters themselves. To provide insight into how the standard time per task can be determined in a similar environment, a short explanation of the method used is given the trolley preparation task. As recommended by Groovers [27], we add 5% allowance time to each standard time.

Trolley Prep

This operator prepares the trolley for the trolley picker. For each trolley the operator needs to take an empty trolley and place coloured paper sheets at the side of the trolley, which indicate the priority. For the remainder, the workload of the trolley prep is dependent on the number of boxes that need to be placed on the trolley. To determine the process time to prepare a trolley as accurately as possible, formula 4.4 is used:

$$Prep\ time\ Trolley = Fixed\ time + Number\ of\ boxes * Variable\ time \tag{4.4}$$

The fixed time represents the time to take the trolley and put the coloured paper at its side and the variable time represents the time it takes to place one box on the trolley.

Table 4.2: Standard times per Task, confidential information is hidden

TASK	DRIVER	AVERAGE STANDARD TIME (SEC)	BASED ON
TROLLEY PREP	Trolley + Boxes	*Confidential*	Direct time studies
TROLLEY PICK	Trolley	*Confidential*	Data
TROLLEY DROP	Trolley + Boxes	*Confidential*	Direct time studies
BOX FILLING	Boxes	*Confidential*	Direct time studies
DOCUMENT APPLY	Boxes	*Confidential*	Direct time studies
PALLETIZING	Boxes	*Confidential*	Direct time studies
CONSOLIDATION	Boxes	*Confidential*	Direct time studies
NESTING	Boxes	*Confidential*	Direct time studies
HOPT PICK	Boxes	*Confidential*	Data
REACH PICK	Pallet	*Confidential*	Data
EPT PICK	Trolley	*Confidential*	Data
PALLET	Pallet	*Confidential*	Direct time studies

Trolley Pick

In addition, the WMS system stores the timestamps of some process tasks. For example, the first and last moment that a trolley picker has scanned the trolley are again saved in the WMS system as barcode scanning is required to confirm the pick. It should be noted, however, that we must be extremely careful with drawing conclusions from the data. The picker can also perform other tasks between the first and last pick, such as cleaning activities, reporting non-conformities to a manager, helping new colleagues or taking a break. Groovers [27] also warns of this when using historical data.

For Ceva it is important to get an overview of the standard process times per task, according to the current task description. The book of Groover suggests dividing tasks into different work elements, which can be measured in direct time studies. Especially for the task of trolley operator, this is very challenging as a lot of factors influence the standard time. The standard time is influenced by the number of unique locations, the picks per location and the walking distance. At this moment, the supervisor expects that there exists a linear relationship between the number of pick lines and the processing time on a trolley. The historical process time of a trolley operator is stored per trolley in the WMS system SAP. We have tested the linear relationship between the number of pick lines per trolley and the process time to process the trolley for a historical data set of 3.5 months. In Appendix D, the linear regression formula and box plot diagram is plotted against the data. Even though the process time of the trolley increases with the number of pick lines, the R-value of the linear regression formula is only 0.33. This is because the process time varies per trolley.

In addition we tested for a linear relationship between the process time per trolley and the number of boxes, which led to a really low R-value of only 0.09. The linear regression plot and box plot are shown in Appendix D. Also the boxplot shows that in fact the average process time per trolley only differs for a small number of boxes (1-6), after that the average process time becomes flat. Also the spread of the process times per box is very large.

Lastly, we tested the linear relationship between the process time per trolley and the number of unique locations that need to be visited. The linear regression plot and box plot of this relationship are also depicted in Appendix D. The most optimal Linear regression formula leads to a R-Value of 0.41. This means that for 41% of the trolleys, the process time can be deduced based on the number of locations that have to be visited within the warehouse. This is an improvement compared to predicting the process time based on pick lines, but it is still not a reliable method. The reason that only the number of unique locations does not give a good indication of the process time is because it says little about the walking distance.

Currently, no distance matrix is available to determine the walking distance between the locations. Most likely, it will help to determine the process time of a trolley more accurately. However, it is only of added value for known pick lines, as for unknown pick lines the picking locations are of course still unknown.
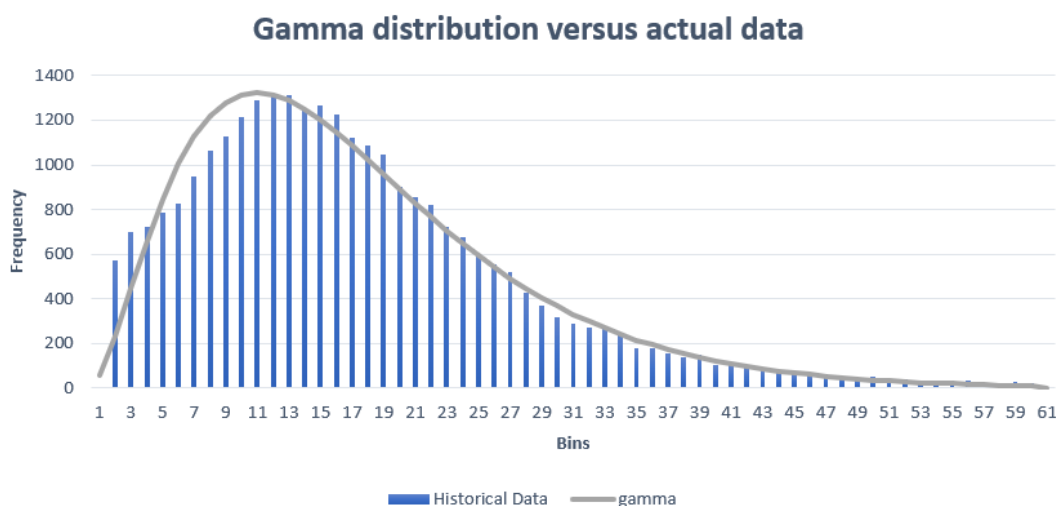

Figure 4.6: Gamma distribution fitted to the historical process time of trolley picking

Therefore we tested if an empirical distribution fits the historical data, such that the process time can be simulated by drawing it from the distribution. We tested multiple distributions such as gamma, Weibull, Poisson and lognormal. The gamma distribution, as depicted in Figure 4.6 fits the historical data the best. However, no empirical distribution passed the statistical chi square test.

We decided therefore to randomly extract a process time from the historical data for the process times of trolley pick and HOPT pick. A standard time per trolley deviates too far from reality, as in practice the trolley process time deviates between one and sixty minutes. For the remaining tasks, the standard time as mentioned in Table 4.2 is used.

4.3.2 Initial workforce scheduling solution

As mentioned, after the standard times per task are known, the initial workforce scheduling needs to be determined. To give an impression of how an initial solution looks, an example is given in Table 4.3.

Table 4.3: Example of an Initial workforce schedule solution

TASK	# FTE MORNING SHIFT	# FTE EVENING SHIFT
TROLLEY PREP	1	1
TROLLEY PICK	7	7
TROLLEY DROP	1	1
BOX FILLING	1	1
DOCUMENT APPLY	1	1
PALLETIZING	2	2
CONSOLIDATION	1	1
NESTING	1	1
HOPT PICK	1	1
REACH PICK	2	2
EPT PICK	1	1
PALLET	2	2

To determine the initial solution, we assume that tasks are not consecutive. Based on the known pick lines, the point forecasts and the standard time per tasks, the total process time per tasks can be calculated. Formula 4.5 shows how the number of trolley picker FTE that is required to process the total demand is calculated.

$$Number\ of\ Trolley\ Pickers = \frac{Total\ Trolley\ Pick\ time}{Hours\ available\ per\ shift} \tag{4.5}$$

This is the absolute minimum number of trolley pickers needed to process demand. It is important to use an initial solution that does not have any FTE overcapacity, therefore the workforce schedule model only needs to add FTE to the planning. If the model starts with more FTEs than necessary, this cannot be scaled down anymore. This procedure is done for every volume dependent task.

4.3.3 Workforce Schedule Model

Once the initial workforce schedule model is determined, all the required input data is available to run the workforce schedule in Excel. The pseudo code of this tool is shown in Table 4.4. Via a Power Query connection in Excel, all the up-to-date information from the SAP WMS system is loaded into Excel. This information contains all the known pick lines. The first step is to filter only the data applicable for this department. For most pick lines the TrolleyID and BoxID are already generated and those can be read from the data. However, for the pick lines that have already arrived but are not grouped to a TrolleyID by the waveplanner yet, those IDs are not known yet. In order to be able to include these pick lines in the scheduling tool, two discrete distributions have been made based on historical data. The first shows how the number of pick lines belonging to a BoxID are distributed and the second shows how the number of

boxes belonging to a TrolleyID are distributed. Both distributions are used to generate the BoxID and TrolleyID if needed. The process of determining these distributions is explained in detail in Appendix E.

In section 2.5.1, six different workflows are distinguished within the department. Each workflow represents a sequence of tasks that have to be processed to finish the pick line. Each pick line has to be processed according to one of these workflows. Below, all workflows are briefly described again, including the tasks that belong to the workflows.

1. Standard Trolley Pick: *Trolley Prep->Trolley Pick ->Trolley Drop->Box Filling->Document Apply-> Palletizing*
2. Trolley Pick Consolidation: *Trolley Prep->Trolley Pick ->Trolley Drop->Box Filling->Document Apply-> -> Consolidation-> Palletizing*
3. Trolley Pick Nesting: *Trolley Prep->Trolley Pick -> Nesting*
4. HOPT Pick: *Trolley Prep-> HOPT Pick ->Trolley Drop->Box Filling->Document Apply-> Palletizing*
5. Reach Pick: *Reach Pick-> Pallet*
6. EPT Pick: *EPT Pick -> Pallet*

Table 4.4: Pseudo Code Workforce Scheduling Tool

Algorithm: Pseudo Code Workforce Schedule
1. INPUT: Known Pick Lines
2. INPUT: Forecast output, expected number Unknown Pick Lines
3. INPUT: Historical Unknown Pick Lines, used to simulate
4. INPUT: Standard times
INPUT: Initial workforce schedule
5.
6. Filter the applicable Known Pick Lines for this department
7. Determine the unique Trolley IDs (solely workflow 1,2,3,4)
8. Determine the unique Box IDs per workflow
9.   In case TrolleyID or BoxID is unknown, use the discrete distributions as explained in Appendix E.
10. Randomly select historical Unknown Pick Lines to fill the gap between Forecasted pick lines And the Known Pick Lines
11.
12. Set model parameters to the initial Workforce Schedule
13. *Do until stopping criteria is met*
14. Run the FJSP heuristic (Explained in section 4.3.4)
15. Determine the shift and task that is the largest bottleneck within the workflow
16. Increase one operator in the bottleneck shift and tasks
17.
18. Output: Workforce schedule, expected backorders, expected end shift time

For all known pick lines, all the information is available to determine according to which workflow they are processed. For the unknown pick lines, this information is not available. Due to the point forecasts, it is known how many unknown lines per MHE type are expected to arrive the next business day. To gather all the input information for the unknown pick lines, we randomly draw the pick lines from historical unknown pick lines. Hence, the point forecast indicates how many Trolley, HOPT, Reach and EPT pick lines should be drawn from the historical data.

The point forecast of Trolley picks makes no distinction between workflow 1,2 and 3 as they are all considered trolley picks. Therefore an analysis has been performed in Appendix F, to determine the ratio between standard trolley pick lines, consolidation pick lines and nesting pick lines. Since this ratio remains almost constant, we assume that by selecting random trolley lines, this ratio remains intact. The arrival and

cut-off times of these randomly selected historical lines can be used directly for the workforce schedule model. The arrival date and due date has to be changed to the next business day.

We then set the number of available operators per task and per shift equal to the initial workforce schedule, after which an estimation of backorders and expected end shift time can be made using the FJSP model. This is explained further in section 4.3.4.

4.3.4 Flexible Job-Shop Schedule solved with heuristics

The FJSP model is applied according to the flowchart depicted in Figure 4.7. For all pick lines that are processed according to workflow 1 to 4, it is first determined which unique trolley batches exist. These trolleys are sorted by due date and cut-off time, so according to the EDD priority dispatch rule. According to this sorted order, each trolley is matched with the earliest available trolley prep operator. This dispatch rule has been chosen as in practice the trolleys are also processed on the basis of due date. The trolley that has to be finished first is also processed first.
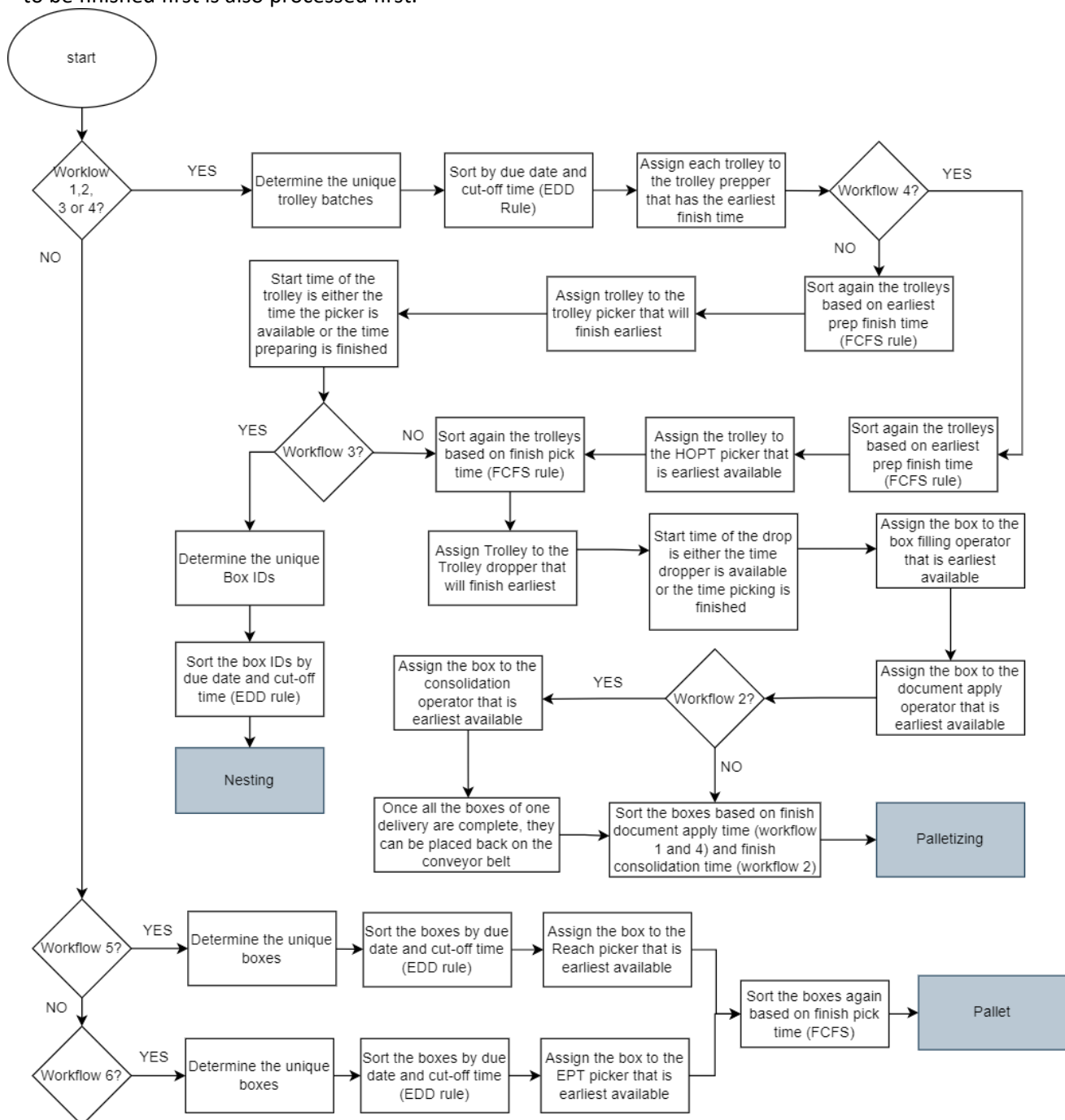

Figure 4.7: Flowchart of the Flexible Job-Shop Schedule with the EDD and FCFS dispatch rules

For workflow 1 to 3, the trolleys are sorted again ascending on the preparation finished time. Hence, the first come first served dispatch rule is used. Thereafter, for each trolley, the first available trolley picker is then selected to pick all the items. The start time of the trolley picker will either be equal to the preparation finished time or the first time a trolley picker is available. This depends on which event happens the latest as both the trolley needs to be finished prepping as well as a trolley picker needs to be available before the trolley pick can start. For workflow 4, exactly the same logic is followed however a HOPT picker is assigned to the TrolleyID instead of a trolley picker.

Workflow 3, is directly processed by nesting after the picking is finished. Nesting prioritises based on due date. However it does need to take arrival times into account.

Workflow 1, 2 and 4 is sorted again based on pick time. Each TrolleyID within these flows are assigned to the next available trolley dropper. The start time of the trolley drop will again either be equal to the pick finished time or the first time a trolley dropper is available. After the boxes are dropped to the conveyor belt, the boxes no longer need to be sorted for the box filling and document applying stage as both workstations process the boxes in the same order.

After the document apply, in case the boxes below to workflow 2 they go off to the consolidation workstation. Here the boxes are temporarily stored until all the boxes that belong to the same delivery are complete. Then they are placed back on the conveyor belt, where they meet the boxes from workflow 1 and 4 again. The BoxIDs are sorted one last time to determine the time that they arrive at the palletizing workstation. For the boxes that belong to workflow 1 and 4, the arrival time equals the finish time at the box filling workstation. While for the boxes belonging to workflow 2, the arrival time equals the finish time at the consolidation workstation. The boxes are finally processed by a palletizing operator after which they are completed.

All the boxes that belong to workflow 5 are sorted according to the EDD priority dispatch rule and then processed by the first available reach operator. The boxes that belong to workflow 6 are similarly sorted according to the EDD priority dispatch rule and then processed by the first available EPT operator. Then the boxes are once more sorted based on either the finished EPT time or the finished Reach time and are processed according to FCFS by the pallet operator.

Once this process is done for all boxes, we can determine the expected backorders and finish shift time. This is explained in section 4.3.5.

4.3.5 Determine the bottleneck and update the initial solution

In the FJSP model, the shift end times are entered as hard parameters. In spite of this, it is allowed to start with the next box or trolley before the end of a shift and to finish it after the end time of the shift. Once the end time has passed, it is no longer allowed to start with a new order. For each workstation, a record is kept of how many trolleys/boxes could not be processed. Hence, the expected backorders with the scheduled workforce capacity.

In addition, the time at which he completed his last work is stored for each operator. If there is no more work, this is before the normal end time of the shift.

In the first iteration, there will be backorders at some workstations as the workforce capacity is scheduled based on a utilisation of 100%. As the pick lines have to be processed by a fixed sequence of tasks, especially the tasks at the end of the workflow cannot process the orders on time.

In the next iteration, we add one operator to the workstation, with backorders, that is earliest in the workflow sequence. Thus, in case the trolley prepare workstation is not capable of processing all boxes on time, we add one operator to that workstation. We have chosen this method because all workstations that come later in the workflow will suffer from, in this case, the trolley prepare workstation.

We keep adding operators to the original workforce schedule, till all workstations are capable of processing all pick lines on time. The workforce schedule that is entered in the FJSP at the latest iteration is the most optimal workforce schedule.

4.4 Summary

To answer the main research question, first ten forecasting models from literature are implemented and evaluated to determine which models are best suited to predict the demand for each of the four MHE types. The literature study has shown that within the statistical forecast models, the ETS model and the ARIMA models can potentially yield good results. From the preliminary analysis, both weekly and quarterly seasonality can be distinguished for trolley pick lines. Therefore the ARIMA model is implemented twice, once for the weekly seasonality which is referred to as ARIMA1 and once for the quarterly seasonality which is referred to as ARIMA2. In addition, the performance of a combination forecast is evaluated which combines the Naïve, ETS, ARIMA1 and ARIMA2 model.

These four individual models as well as the combination model can be used to forecast the Trolley, HOPT, EPT and Reach picks based on the historical pick lines of each type. In addition, the same models can also be used in combination with the top-down approach. This approach forecasts the aggregated pick lines and divides the pick lines over the four MHE types based on the ratio.

In addition, at the scheduling moment a fraction of the total lines are known. Via the linear regression model, we will determine the linear relationship between the dependent variable which are the known pick lines and the independent variable which are the total pick lines per MHE type. This relationship is used to predict the total pick lines based on the known pick lines for the validation data set. The objective is to evaluate all these introduced forecasting methods and select the most accurate model per MHE type. Hence, this can either be one of the statistical forecast models or the linear regression model.

Secondly, this chapter described how the standard times per task are determined. Two methods have been used for this, namely direct time studies and historical data. For four out of the twelve tasks, the WMS system saves the process time. This leads to much more measurements than can be collected in a short period of time through time studies. For the remaining eight tasks, the WMS system does not store any processing times. Therefore we determined via direct time studies the standard times.

Lastly, the decision-support workforce scheduling model is introduced in this chapter. Based on the known pick lines, the output of the forecasting model and the standard times per task, this will determine the most optimal workforce schedule. This is done via an iterative model that calculates the expected backorders and finish shift time via the FJSP model. As long as there are still expected backorders, the workforce schedule will increase the bottleneck task by one. The new workforce schedule is entered into the FJSP model, which calculates the objectives again. Once the outcome of the FJSP expects no backorders, the optimal workforce schedule is found.

5 Model Results and Validation

The previous chapter introduced ten forecast models and the workforce scheduling model. Section 5.1 elaborates on the results of each model and selects the best performing model per MHE type. Section 5.2 displays the results of the workforce scheduling tool and describes the validation of the tool. Lastly, all results will be concluded in section 5.3.

5.1 Forecast results and validation

For each of the four MHE types, ten different forecast models are applied to predict the number of pick lines for the next business day. These forecast models and their implementation methods are introduced in section 4.1. This section answers the sub-research questions: *Which validation model is selected to assess and validate the forecasting model?* and *How well does the forecast model perform?*

The first five models evaluated are the Naïve model, the ETS model, ARIMA1 with weekly seasonality, ARIMA2 with weekly and quarterly seasonality and a combination of these four models. We then apply the ETS, ARIMA2 , ARIMA2 and the combination model again for the aggregated time series in combination with the Top-Down approach and finally linear regression is applied. To be able to compare the forecasting error of different MHE time series, it is useful to display the forecasting error as a percentage of the time series value. The selected measure for the forecast error is MAPE. Figure 5.1 depicts the MAPE values of each of the forecast models for EPT, HOPT, REACH and Trolley time series. The models in which the aggregated time series have been used to subsequently apply the Top-Down approach are indicated in Figure 5.1 by AGG, which is an abbreviation for aggregated.
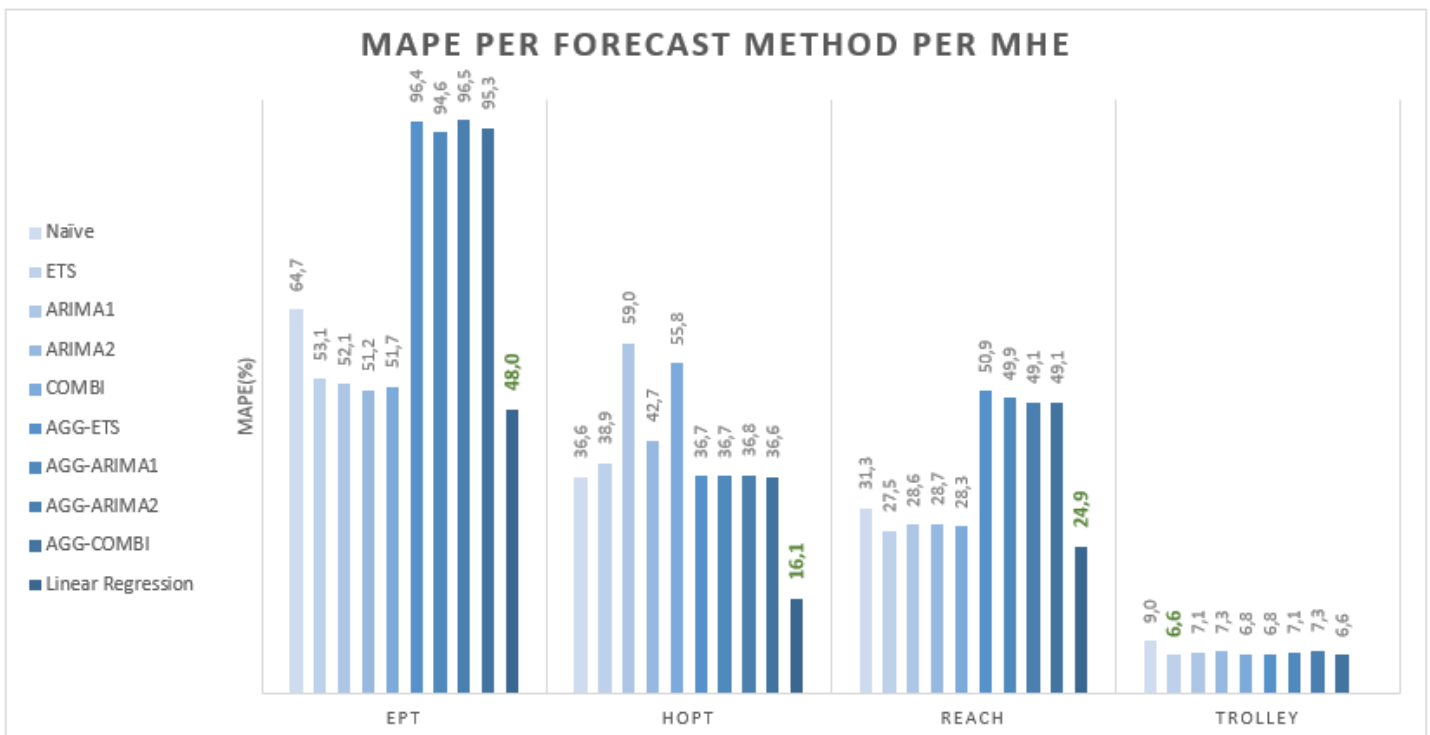

Figure 5.1: MAPE for all evaluated forecast models per MHE

The Naïve model is applied to compare the MAPE values. MAPE values are not scaled errors, so the MAPE values of the other models must be compared with a simpler model such as the Naïve model. This way, the MAPE value can be put into perspective. If the MAPE value of the Naïve forecast is lower than that of more complex models, it is not worth the time investment to implement and use these complex models.

In the remainder of the section, all the forecast models and their performance per MHE type are discussed in detail. In section 5.1.1 the results of ETS and both ARIMA models are described. Section 5.1.2 focuses on

the combination forecast and shows the determination of the compound weights. Section 5.1.3 describes the results of the Top-Down approach and section 5.1.4 shows the results of the linear regression model.

5.1.1 Results of ETS, ARIMA1 and ARIMA2

Figure 5.1 shows that the ETS model and both ARIMA models result in a smaller forecast error, expressed in the MAPE value, than the Naïve forecast model. This applies to all MHE types. It can be deduced from this that all three models lead to a more accurate forecast compared to the Naïve model. Hence, the models have potential and are therefore further taken into consideration.

While the ARIMA models vary in performance, as for Trolley and REACH, ARIMA1 is more accurate while for EPT and HOPT pick, ARIMA2 is more accurate, the ETS model outperformed both ARIMA models for all MHE types. We focus on the ETS model used for Trolley Picks, which achieves a MAPE value of 6.6%. For Trolley picks this MAPE value is the most accurate out of all evaluated models. Hence, this model is selected to forecast the trolley picks.

As described in section 4.1, based on the training data set, the model's parameters are selected. For the trolley pick lines, an additive damped trend component and a multiplicative seasonal component are selected via the ETS function is R. The combination of these parameters within the ETS model is in fact the *Holt-Winters damped method* which has already been described in section 3.3.1. The forecasted values of Holt-Winters is calculated with formula 4.1 [44], where l_t represents the time series level at time t, b_t the slope at time t and s_t the seasonal component of the series at time t. As the time series seasonal frequency has a limit of 24 seasons, only weekly seasonality is taken into account. Therefore m, which denotes the number of seasons, is equal to five. The level, scope and seasonal component are calculated respectively via formula 4.2, 4.3 and 4.4.

$$\hat{y}_{t+h|t} = (l_t + \phi b_t)s_{t-m+h_m^+} \qquad 4.1$$

$$l_t = \alpha\left(\frac{y_t}{s_{t-m}}\right) - (1-\alpha)(l_{t-1} + \phi b_{t-1}) \qquad 4.2$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)(\phi b_{t-1}) \qquad 4.3$$

$$s_t = \gamma(\frac{y_t}{l_{t-1}} - \phi b_{t-1})) + (1-\gamma)s_{t-m} \qquad 4.4$$

Based on the test data set, the smoothing parameters α, β, γ and ϕ are optimised in R. This resulted respectively in the values 0.49, 0.001, 0.001 and 0.97. These smoothing parameters are considered fixed for the validation. The performance of the Holt-Winters damped model is determined by the validation data set, unknown data that the model has not seen before, to prevent overfitting. As the time series cross validation prescribed, predictions are always made one day ahead. After that, the actual number of trolley pick lines of that day is assumed to be known, with which formulas 4.1 to 4.4 are recalculated. Subsequently, the next day can be predicted. This procedure is repeated for each of the 55 data points in the validation set. The MAPE value is then calculated over the whole validation set to give the best possible impression of the model's performance.

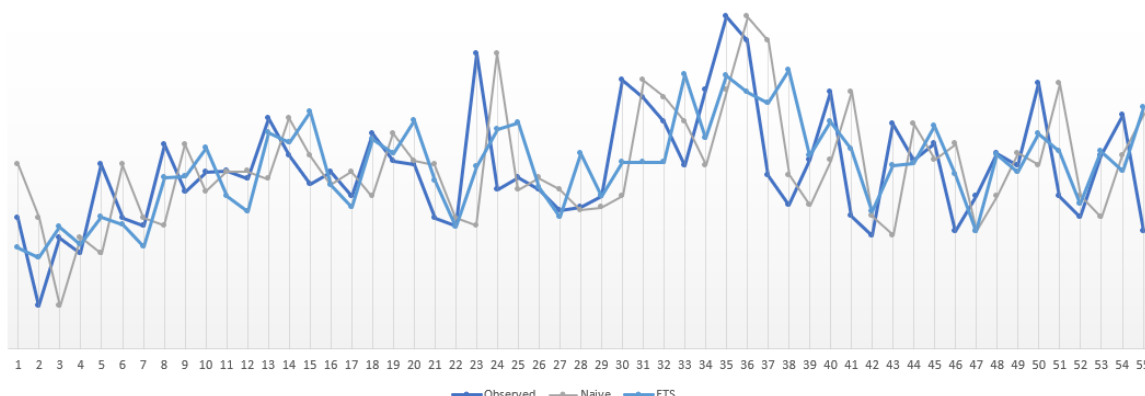Actual Trolley pick lines versus the forecasted Trolley pick lines

Figure 5.2: Actual Trolley pick lines versus the forecasted trolley pick lines by the Naïve and ETS model

Figure 5.2 shows the actual Trolley pick lines in the dark blue for each day within the validation data set. The grey line depicts the Naïve model and the ETS model is shown as the light blue line. The Y-Axis has been omitted so that no confidential information can be derived from the graph. From Figure 5.2 it can be deduced that the ETS model is able to adjust the level quickly as the actual number of lines increases or decreases, this is due to the relatively high value of the alpha smoothing parameter.

5.1.2 Results of the combination model

As mentioned by Hyndman [4] combining multiple forecast models often leads to smaller forecast errors than any of the compound models can achieve individually. As mentioned in section 4.1.7, the test data set is used to determine the optimal compound weights of each of the forecast models. The combination model evaluated in this research combines the Naïve, ETS, ARIMA 1 and ARIMA2 model. Figure 5.3 shows how the compound weights of the four models change within each iteration within the test data set for Trolley picks. For the first iteration, each model has an equal compound weight of 1/4. Based on the MAPE values of the individual models, the compound weights are adjusted to minimise the MAPE value of the combination forecast. The weight of the Naïve model drops quickly to zero. The other compound weights at the end of the test data set are 0.27, 0.20 and 0.53 for respectively the ETS, ARIMA1 and ARIMA2 models. These compound weights are assumed fixed when the model validation is performed over the validation test set.
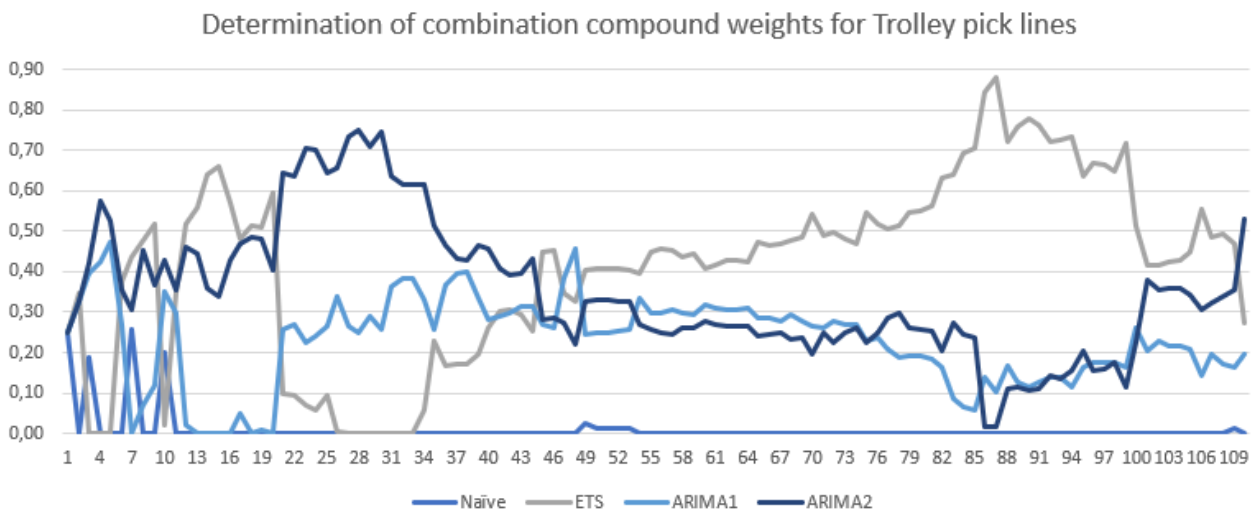

Figure 5.3: Determination of compound weights of the combination forecast model

Even though the MAPE value of the combination model approaches the MAPE value of the best performing compound model, for the EPT and Trolley pick lines, the combination model does not outperform any of the compound models. On this basis, this model is not applied further within this research.

5.1.3 Results of the Top-Down Approach

For the Top-Down Approach, the historical aggregated pick lines are used as input instead of the pick lines of the individual MHE types. Subsequently, the EPT, ARIMA1 and ARIMA2 forecast models are again used to predict the number of aggregated pick lines. The total number of predicted lines is split across the MHE types. Based on the training and test data set, the calculated shares of respectively EPT, HOPT, REACH and Trolleys picks are 2.6%, 0.8%, 1.8% and 94.8%. The aggregated forecast is split over the MHE types by these percentages.

Especially for HOPT and REACH picks, this method leads to a poor forecast accuracy. Figure 5.1 shows that the MAPE values for HOPT and REACH picks are significantly higher than the MAPE values of the Naïve model. For EPT and Trolley picks, the results of the Top-Down approach beats the Naïve model and approaches the MAPE value of the ETS and ARIMA models. However, as none of the models that use the

aggregated time series as input outperforms the ETS model, it is not advised to use this method for this application.

5.1.4 Results of the linear regression model

As mentioned in section 4.1.9, the linear regression uses a different approach than the nine forecast models discussed so far in this chapter. Instead of predicting the daily pick lines based on historical data, the linear regression model tries to find a linear relationship between the number of known pick lines and the number of daily pick lines. As by the time the forecast model has to run, thousands of trolley pick lines are already known, the possibility of a relationship between the two variables cannot be ruled out in advance. Since the number of daily pick lines must be predicted with the number of known lines, the number of known lines represents the independent variable and the number of pick lines represents the dependent variable. To check the relationship per MHE type, the *lm* formula within R is used.

How well the formula describes the relationship can be deduced from the R-value and the P-value. The P-Value shows how likely it is that the independent variables are valuable to predict the dependent variable. In general, it is assumed that with a P-value smaller than 0.05, then the independent value is a valuable addition to the model. With a P-value higher than 0.05, it is better not to include that independent variable in the model. The R-value indicates the goodness of fit of a model. Models that fit the data well, obtain an R-Value close to one, while models that perform poorly have a value near zero.

The R-value of the linear regression formula found for Trolley picks is 0.01942. This means that the model is able to explain 1,9% of the data variability. From this it may therefore be concluded that the model is not capable of predicting the Trolley pick lines. The linear regression model for the aggregate pick lines results in a similar R-Value and P-Value. The R-Values for EPT, HOPT and Reach picks are respectively 0.63, 0.78 and 0.54. The P-Values of each model is below 2.2e-16 for each MHE type, which indicates that linear regression could work well for these three MHE types. The linear regression formula that results in the lowest 'least squared error' is shown in Table 5.1 per MHE type. To give a better impression of how well a linear regression describes the relationship between the variables, the linear regression formula for HOPT picks is plotted in blue in Figure 5.4. The black dots represent the training data used to determine the relationship.

Table 5.1: Linear regression formulas for the MHE types EPT, HOPT and Reach

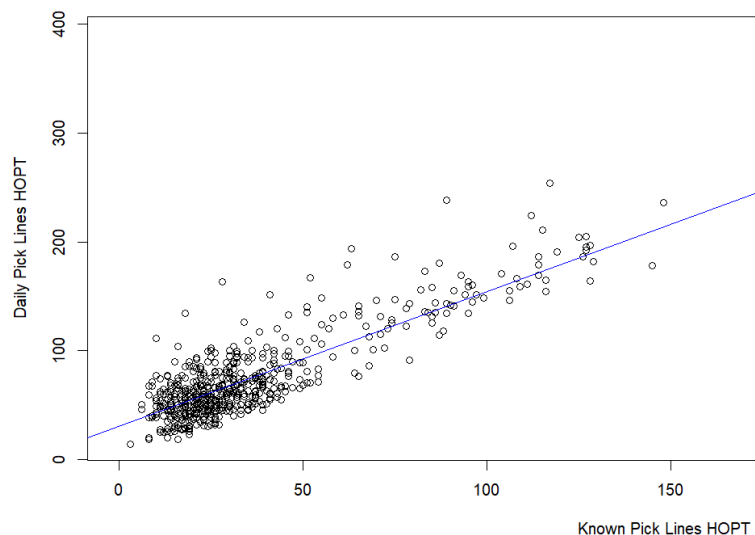| MHE TYPE	LINEAR REGRESSION FORMULA
EPT	$Pick\ lines = 90.58 + 0.95 * Known\ Lines$
HOPT	$Pick\ lines = 30.55 + 1.25 * Known\ Lines$
REACH	$Pick\ lines = 59.23 + 1.15 * Known\ Lines$

Figure 5.4: Linear regression formula HOPT, Daily pick lines versus Known lines

As shown in Figure 5.1, the linear regression model well outperforms any of the traditional statistical forecast models for the EPT, HOPT and Reach picks. Hence, this forecast model is selected for those three MHE types to predict the next day pick lines.

5.1.5 Interpretation of the MAPE value

A MAPE value alone does not provide a complete picture of the performance of the forecast model, as it is an average value. To give a more detailed impression of what an average MAPE value of, for example 6.6% means, Figure 5.5 depicts the Box Plot diagram of the MAPE values of the best performing models per MHE type. The figure shows that trolley picks can be predicted reasonably accurately with an average forecast deviation of 6.6%. The box plot shows that 75% of the point forecasts for trolley picks deviate between 2.1% and 9.5% from the actual pick lines, which is a good basis for the workforce schedule model. It is significantly more difficult to forecast the EPT, HOPT and Reach lines, as the MAPE values are larger. However, as the volumes are so much lower, maintaining a higher safety margin will not affect the required number of FTEs that much compared to the Trolley picks. For example, on average the number of HOPT lines is 77, when the forecast deviates by 30% from reality that leads to an increase or decrease of 23 lines. By scheduling an operator that is cross trained, it is possible to deal with a higher deviation in EPT, HOPT and Reach pick lines.
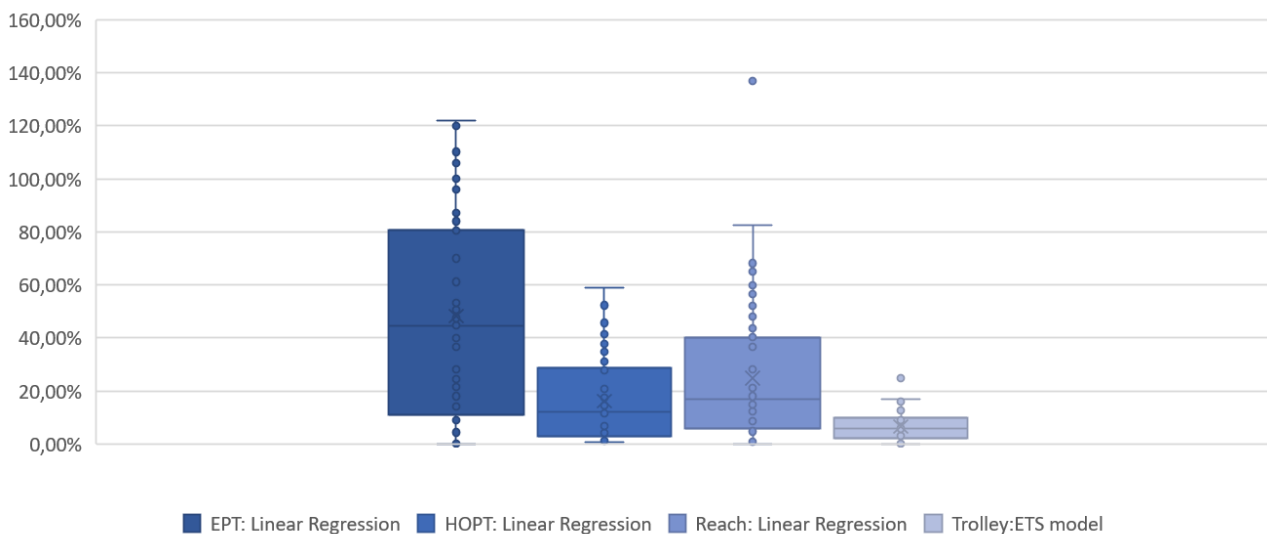

Figure 5.5: Box plot of the MAPE value of the selected models for each MHE type

5.2 Workforce schedule results and validation

This section answers the sub-research question: *How is the scheduling tool validated?*

We programmed a macro in Excel which displays the results of the workforce schedule in the form of a Gannt chart as shown in Figure 5.6. There are two buttons in the excel file, at the right upper corner. The first is used to refresh the query, this ensures that all data is retrieved from the WMS system again and that the data is up-to-date. The second button invokes the workforce scheduling tool model. Before the supervisor uses this button, he needs to enter the number of operators that are on the current shift. This is built in as an additional check, so that sick employees are not included in the planning.

After the supervisor presses the buttons, the optimal workforce schedule is generated as described in section 4.3. Each row in the Gannt chart represents one operator. Which type of operator that is, and for which shift he is scheduled is stated on the left. At the top centre of the Gannt chart are the hours of the day displayed. The schedule starts at 3PM, as that is when the supervisor creates the schedule. The shift in operation works until 9PM. The hours 6AM to 21PM at the right-hand side represent the next business day.

Each blue rectangular of the Gannt chart represents a box or a Trolley. The width of the blocks indicates the process time of the trolley or box. If the blocks do not connect directly to each other, this means that the relevant operator has to wait for the next box or trolley. This is of course something we want to avoid as much as possible.

In the fourth column in the Gantt chart, as shown Figure 5.6, the supervisor also receives an overview of the total idle time per employee. An example of a task that consists of a lot of idle time is the consolidation task. As shown in Figure 5.6, the consolidation operator's are waiting for boxes for more than half their shifts. This is an indication that the supervisor can ask the consolidation operator to support other operators during his shift. The idle time parameter will also be of great importance later in the control room to reallocate personnel optimally.

In addition, the time at which every operator processes the last box/trolley according to schedule can be easily read from the Gantt chart. For example, the eleven trolley operator's that are scheduled for the evening shifts are expected to finish all the orders of that day by 18:30 PM. As a result, these operators are available to process additional demand between 18:30 PM and 21:00 PM. Hence, if more demand arrives than forecasted, there is time to process these lines to prevent backorders. In case all orders with today's PGI date are processed, the operators can start processing next day's orders.

Currently, supervisors schedule the same number of operators in the morning shift as in the evening shift. Ceva values it highly that the decision making is done by humans instead of the model. Therefore, we created this workforce schedule as a decision-support tool. To give an example, the HOPT picker that is scheduled for tomorrow's evening shift has according to the schedule nothing to process. The Gantt chart will help to visualise these schedule improvements. Based on this, the supervisor can refine the planning at his own discretion.

We presented the Gantt chart to the supervisors, to receive their opinion about the tool. Their feedback was that this tool can support them well in making a schedule because they can extract a lot of relevant data from the Gantt chart. The standard times have also been validated by the supervisors, as a result we know that the standard times are representative of practice.

The validation of the workforce scheduling model is based on the hours report. This is a daily overview on how many hours have been clocked on each task. The number of hours deployed per task has been converted into the number of FTE deployed per task. These parameters are entered into the workforce scheduling tool along with the pick lines for that day. The outcome of the workforce scheduling model shows whether all lines are processed on time. By comparing the number of expected backorders with the actual number of backorders, it is validated whether the schedule of the workforce scheduling tool is accurate.
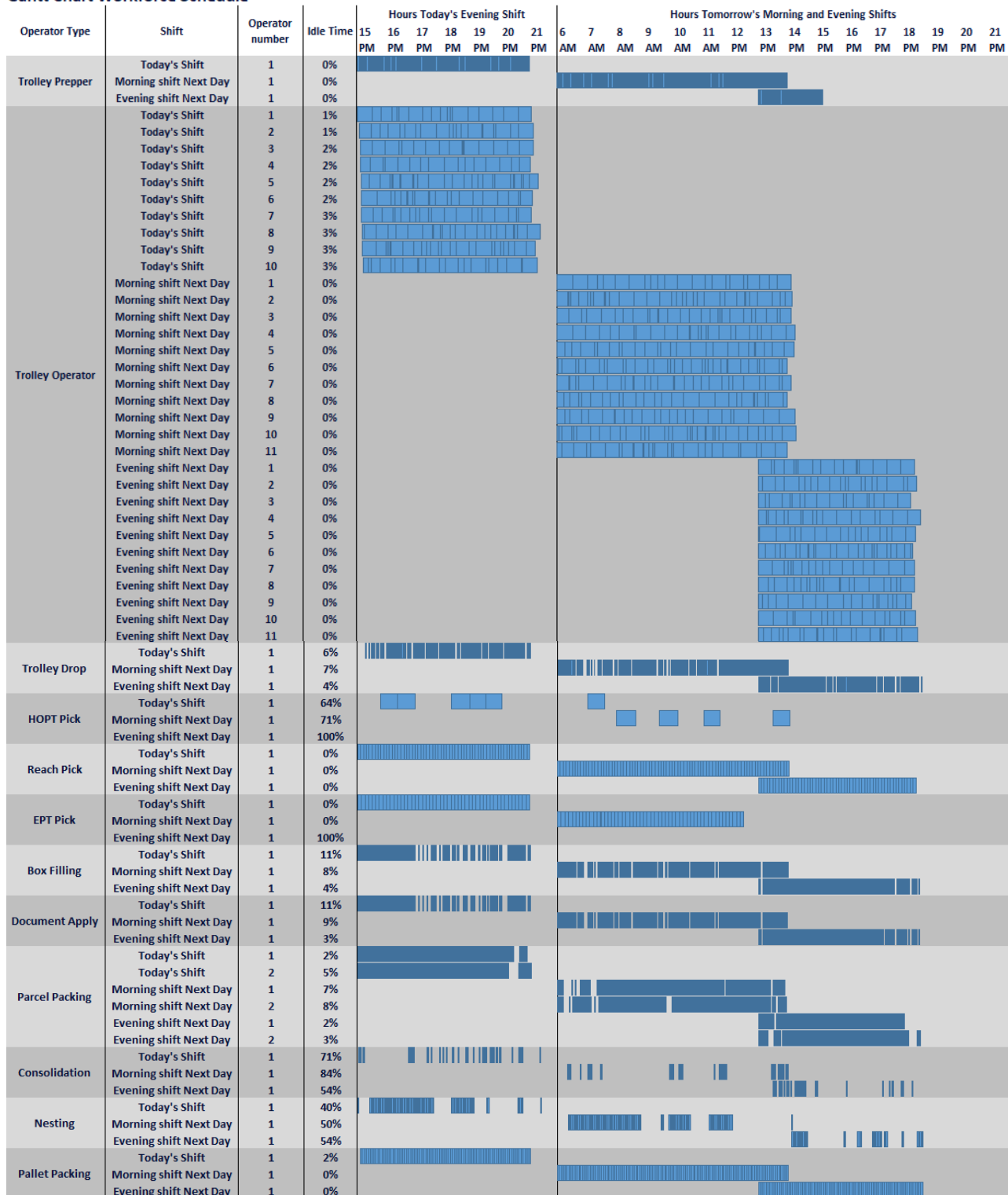
Gantt Chart Workforce Schedule

Figure 5.6: Dashboard of the workforce scheduling tool

5.3 Conclusion

After evaluating all forecast models, it appears that the trolley pick lines can be forecasted the most accurately by far. The Holt-Winters variant of the ETS model even achieves a MAPE value of 6.6% on average, measured over a period of 55 business days. Applying the Naïve forecast Ceva used before over the same data set leads to a MAPE error of 11.3%. Hence, a decrease in the mean error percentage of 41%. A more accurate forecast is an important first step to make a better workforce schedule.

For the other three forecast models it is more difficult to forecast accurately. This is related to the fact that the number of absolute HOPT, EPT and REACH pick lines is quite low, causing a deviation in the forecast to result in a relatively higher margin of error. As the preliminary analysis already shows, there is no seasonal pattern to be found with these three MHE types. With HOPT pick lines, the Naïve forecast even leads to the most accurate forecast compared to all statistical models. This corresponds to Hyndman's theorem [4] that if a time series does not show a seasonal or cyclic pattern, the Naïve forecast model is often the most accurate of the statistical forecast methods.

In addition to the statistical models, we also implemented linear regression. The linear regression model predicts the number of daily pick lines predicted with the number of known lines. Hence, the number of known lines represents the independent variable and the number of pick lines represents the dependent variable. For EPT, HOPT and Reach picks this gives surprisingly good results. The worldwide M3 forecasting competition shows that the static forecast models often still outperform the machine learning models. Surprisingly, this is not the case here. The most logical explanation is that the share of EPT, HOPT and REACH are consequences of storing decisions within Ceva. If products are stored elsewhere a trolley pick can be used instead of HOPT pick for example. Therefore you need the most real time information as possible, which are the known pick lines.

The top-down approach only leads to better results with HOPT pick than the direct forecast of historical HOPT lines. The fact that this does not lead to better results with the other three models is mainly due to the constantly changing ratio in pick lines between the different types. For example, on average 2.6% of the total number of pick lines is an EPT pick. If this is suddenly 3.5%, then the top-down approach is quickly wrong.

We created a decision-support workforce scheduling tool that is capable of determining the required number of FTE based on the standard time and expected demand. Within Ceva's original workforce schedule tool, only four tasks are considered volume dependent. Hence, for only four tasks this model tries to find the optimal number of operators. Within this research we discovered that actually twelve tasks are volume dependent. By including all twelve tasks in the workforce scheduling tool, the number of FTE can be tuned much better to the expected demand. In addition, the Gantt chart provides multiple valuable insights for the supervisors such as the idle time per operator, the expected shift end per operator and the expected backlog. By using this tool, the supervisors will get a better feeling for workforce scheduling, as they will be encouraged to continue making decisions themselves

6 Implementation

Within this thesis, we introduced a forecasting tool in R and a workforce scheduling tool in Excel. As both models are validated, the next step is to embed both models in the daily operation. For this, change management is required as the supervisors need to partly adapt their working way. This chapter gives insight into the operational changes that are required to implement the models and thus answers the sub-research question: *How can the implementation roadmap be optimally designed?* We will first focus on the chances required to implement the forecast tool and sequentially the operational changes to successfully implement the workforce scheduling tool is mentioned.

6.1.1 How should the forecast model be maintained and implemented?

The ETS model is capable of forecasting the number of trolley pick lines the most accurately. To use this model, the realised trolley lines of the day before must be entered daily in an excel input file. Supervisors can extract the realised trolley lines from a report that is sent out daily. Then the model in R can be used to forecast the expected Trolley pick lines of the next day. The computation time is approximately 10 seconds, so the forecast will not be time consuming for the supervisor.

To predict the number of EPT, HOPT and Reach pick lines there is no manual action of the supervisor required. The known pick lines for each MHE type are already saved in a dashboard. This is linked to the scheduling tool such that the pick lines can be predicted automatically via the linear regression formulas. As a result, it takes little time to explain to the supervisors how to use the forecast model.

As Ceva is a dynamic company that undergoes many changes, it is possible that volume ratios change relatively quickly. For example, the ratio between known and unknown pick lines changes or seasonal effects change. Therefore it is recommended to re-evaluate on a frequent basis whether the currently selected models are still the most optimal for each MHE type.

Preferably the re-evaluation is not performed by a supervisor as basic knowledge of R is preferred. This responsibility can better be placed with the new control room team, since they contain the necessary theoretical knowledge. This re-evaluation can always be performed on the basis of the most recent historical data over a period of two years. The same method is used as described in section 4.1, whereby a split is made between training, testing and validation data. Based on the average MAPE values over the validation set, the best model per MHE type is chosen. In case the most accurate forecast model changes for any of the MHE types, the code in R should be adjusted such that the new model and its corresponding hyperparameters are used to forecast the pick lines. The re-evaluation will take about an hour for somebody with the right knowledge and it is recommended to perform this each fiscal quarter. This is also practically feasible since there is an evaluation moment every quarter to determine the KPIs and the performance of the operation.

At Ceva ownership is highly valued. Even though they embrace automatization, it is still important that employees make good judgments. Therefore, the outcome of the scheduling tool is an advice, of the number of employees needed per task. As Ceva is not allowed to schedule part-time operators, there will be no operators scheduled for a few hours only. Based on the idle time per employee and the expected finish time per employee, the supervisor is capable of adjusting the workforce schedule. For example, the supervisor can determine that one operator performs the first four hours of his shift task and the last four hours of his shift task B. As a result, we trigger the supervisor to create the optimal schedule himself, but based on an accurate forecast and current standard times.

To ensure that the workforce scheduling tool remains accurate in the future, direct time studies and data studies are needed to maintain accurate standard times.

7 Conclusion and recommendations

Section 7.1 summarises the main findings of this research and answers the main research question. Section 7.2 describes the recommendations for Ceva Heerlen. Subsequently, section 7.3 provides the research's limitations, its contribution to research and the direction for future research.

7.1 Conclusion

At the moment, Ceva too often exceeds the backorder rate which has been agreed on with the client in a KPI. The consequence of this is, that in addition to a penalty, extra pressure is put on the operation as they have to process the backorders on top of the normal demand. The current scheduling tool is not accurate enough to prevent a backlog situation. This is not only due to the scheduling tool, but also to the lack of a forecast model and the lack of standard times for tasks. Therefore, the research question is:

How can workforce scheduling of operational employees be improved at Ceva Logistics Heerlen by including short-term demand forecasting and the estimation of standard times?

The goal of the forecast tool is to forecast next day's demand as accurately as possible for all four material handling equipment's (*MHE*), which are Trolley picking, HOPT picking, EPT picking and Reach picking. We have evaluated the following ten models: The ETS model, ARIMA1 with weekly seasonality, ARIMA2 with quarterly seasonality, a combination of all the previous. The top-down approach is implemented combined with the ETS, ARIMA1, ARIMA2 and combination model. Lastly, the Naïve model and linear regression were evaluated.

More than 94% of all pick lines are trolley pick lines. Therefore this forecast is the most important, as the largest part of the workforce capacity is ultimately deployed to the tasks that process trolley pick lines. Trolley pick lines show a weekly seasonal pattern, which can be forecasted the most accurately by the ETS model. The optimal model parameters that are selected within the ETS model, make it in fact the Holt-Winters model. An average MAPE value of 6.6% on average is achieved, measured over a period of 55 business days. Compared to the Naïve trolley forecast Ceva currently uses, this reduces the forecast error on average by 4.7%.

For the other three forecast models it is more difficult to forecast accurately. This is related to the fact that the number of absolute HOPT, EPT and REACH pick lines is quite low, causing a deviation in the forecast to result in a relatively higher margin of error. As the preliminary analysis already shows, there is no seasonal pattern to be found with these three MHE types. The forecast model revealed that the best accuracy is achieved by the linear regression model, for the EPT, HOPT and REACH picks. The linear regression model bases the expected number of pick lines at the known pick lines of that respective MHE type. The linear regression formulas as well as the MAPE values are depicted in Table 7.1.

Table 7.1: The most accurate forecast models per MHE type, including the MAPE error

MHE TYPE	BEST FORECAST MODEL	MODEL	MAPE(%)
TROLLEY	ETS	Holt-Winters	6.6
EPT	Linear Regression	$Pick\ lines = 90.58 + 0.95 * Known\ Lines$	48.0
HOPT	Linear Regression	$Pick\ lines = 30.55 + 1.25 * Known\ Lines$	16.1
REACH	Linear Regression	$Pick\ lines = 59.23 + 1.15 * Known\ Lines$	24.9

Based on direct time studies and historical data, we then determined the standard times per task. Together with the forecasted number of pick lines, these values are the input for the workforce scheduling tool.

The workforce scheduling tool generates an initial workforce schedule. Via the Flexible Job-Shop Model, the expected backorders and expected shift ends are determined via two priority rules, the *First Come First Served* and the *Earliest Due Date.* As long as there are expected backorders, an operator is added to the schedule. Every iteration the FJSP model is recalculated until there are no more backorders anymore. The workforce schedule that generates no backorders is intended as advice for the supervisor. This tool will support the supervisor to make a good schedule based on data.

7.2 Recommendations

Based on the evaluation of the forecast tool, we recommend Ceva Heerlen to implement the forecast models as mentioned in Table 7.1. The forecast will certainly help the operation to react more proactively to the demand. It should be noted, however, that the correct selection of the hyperparameters directly influences the accuracy of the forecast. For the daily use of the tool, these parameters are fixed. However as Ceva is a dynamic company that undergoes many changes, it is possible other hyperparameters will become more optimal in the future. Therefore it is recommended to re-evaluate on a frequent basis whether the currently selected models are still the most optimal for each MHE type.

In addition, a good way to deal with a forecast error is to process pick lines ahead. For example, by already processing next day's pick lines during the last two hours of today's shift, there will be extra time to process any unexpected demand tomorrow. If demand turns out not to be higher than expected, supervisors can use this time to again process pick lines ahead. The advantage is that if other unforeseen situations arise, like for example illness, then there is some room to solve this without getting directly into a backlog situation.

To ensure that the workforce scheduling tool remains accurate in the future, it is recommended that direct time studies and data studies are performed quarterly. In case changes are noted, the scheduling parameters should be updated.

7.3 Limitations, scientific contribution and future research

This section describes the research limitations, the scientific contribution and possible future research.

Limitations
The first limitation of this research is that we only have possession over the historical data, to forecast the expected demand. When future events like for example sales can be taken into account, the quality of the forecast improves. However, as this knowledge is not shared with Ceva as a 3PL company, we forecasted purely based on historical data and present knowledge.

A second limitation is that the forecast models cannot be executed on Ceva hardware due to IT limitations. Therefore, the forecast model and the workforce schedule model are not yet embedded in the daily operations. The process has been set in motion to remove these limitations, after which the models will be used daily by the operation.

Scientific contribution
To the best of our knowledge, determining the number of temporary employees required per task at a daily level without knowing the full demand, has not been not done before. To incorporate the unknown demand, predict the total demand via the selected forecast models.

Subsequently, the expected demand needs to be translated to the number of temporary employees per task. Our solution to use a Flexible Job-Shop model to determine an optimal workforce schedule via an iterative approach is therefore new. The major advantage of applying the FJSP model in determining the optimal workforce capacity is that the process sequence is taken into account when making the schedule.

Future research

The models introduced in this research can be extended in multiple ways. First of all, the forecast model, now forecast's only the next day's demand. Within Ceva there is a need to look ahead for a longer period of time, to support strategic and tactical workforce scheduling. Hence, the forecast horizon can be increased.

Additionally, the workforce scheduling tool could be used to run experiments to determine the impact of operational changes. For example, Ceva splits the workforce 50/50 over the morning and evening shifts. The workforce scheduling tool can be extended to determine if this 50/50 split is optimal.

Lastly, priority rules are used to solve the FJSP as realistically as possible. However, research states that priority rules almost never lead to the most efficient schedule. Hence, there is room for optimisation by assigning operators to specific tasks. The workforce scheduling tool can be used to calculate the productivity increase.

Appendices

Appendix A   Description per task

TASKS		DESCRIPTION
PICKING	Technical operator U5	Solves malfunctions of the machine and performs small maintenance
	Packsize	Manually folds boxes, cleans machine on daily basis. Additionally performs small maintenance on the machine
	Trolley prep	Places the boxes on the correct trolley and attaches the coloured priority paper
	Trolley pick	Picks the trolley with highest priority, walks to the location shown on the scanner, scans the products and the box. Subsequently, places the product in the box. After all the products are picked, reports potential abnormalities and drops the trolley at the drop-off location
	HOPT pick	Picks-to-box in the same way as a trolley picks but then at a high-level location in the reserve area. Additionally, picks TO lines and deliver them to parcel packing.
	EPT pick	Picks pallets from ground locations in the reserve area
	C-plein	Relocates pallets from temporary storage location to the correct location within C-locations with pallet jacks
	Reach	Moves pallets to the C-square (C-plein) or to ground locations based on TO lines
	Chasers	Checks if there are products who most likely won't make the cut-off time. Sends a dedicated operator to pick the products and ensures that a parcel packer packs the product. Additionally takes care of the emergency shipments
PACKING	Trolley drop	Moves the boxes of the trolley and place them on the conveyor belt
	Conveyor belt	Performs small maintenance actions. Mentor other operators and support: document apply, box filling and trolley drop.
	Box filling	Adds air sealed bags to boxes to prevent damaged products and checks manually if the right products are within the right boxes.
	Document apply	Applies the right documents to each box
	Consolidation	Temporary stores the boxes in shelves until the order is complete. Then attach the labels to the boxes and places the boxes back on the conveyor belt. Additionally supports palletizing.
	Nesting	Manually places the right boxes on the right pallets
	Palletizing	Scans the boxes, based on the carrier shown in the scanner, the box is places on the right pallet. Subsequently wraps the pallet in plastic when it is stacked till the maximum height.
	Parcel	Chooses the appropriate box size and manually place all the products within the box. Thereafter closes the box and attaches the shipment label.
	Pallet	Packs the whole pallet

Appendix B   Preliminary Analysis for trend and seasonality using time series Pick Lines Trolley

Figure B.1 depicts in black the time series of the total pick lines. The blue line represents the moving average with a time horizon of 30 business days. As explained in section 4.1.2, the weekends are removed from the time series. Additionally, the pick lines at public holidays have been replaced by the average values. Based on the moving average line in Figure B.1, the seasonal pattern of the first year seems to repeat in the second year. The number of pick lines seems relatively stable except for holiday periods, especially around the summer holiday and the Christmas period the pick lines drop.
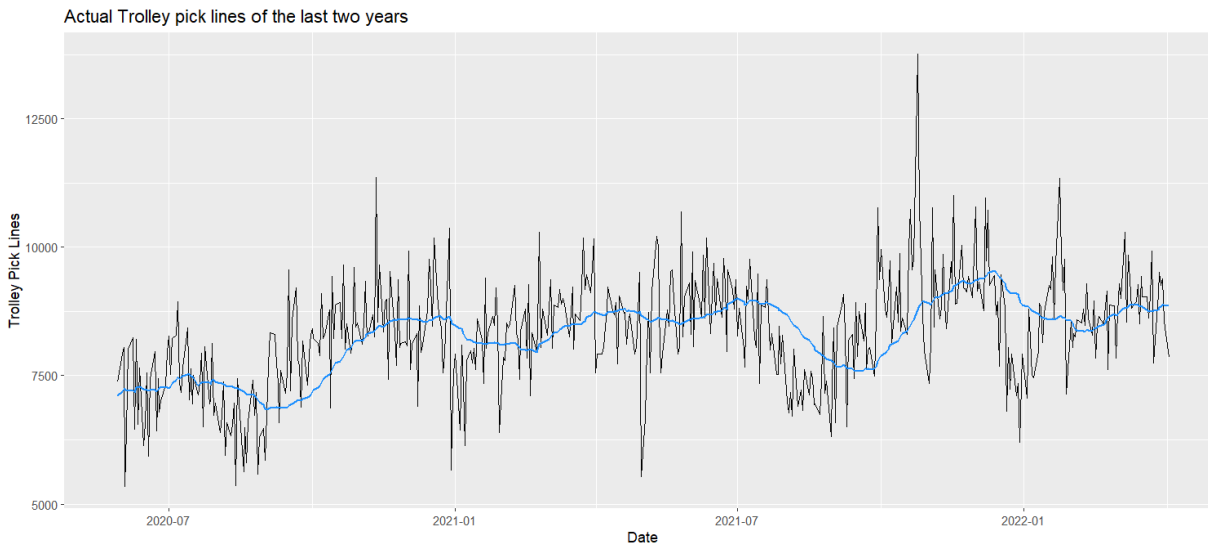
Figure B.1: Actual Trolley pick lines of the last two years. In black the actual pick lines, in blue a moving average with h=30

Experts mentioned that the client operates on a fiscal year basis. Since it is a listed company, they like to close fiscal quarters and years favourably. In the expert's opinion, this often results in sales promotions and therefore a higher workload for Ceva. The time period of the 2022 fiscal quarters are given in Table B.1. Figure B.1 does not show this behaviour as strongly for each quarter, but it does

QUARTER	TIME PERIOD
1ST	01-05-2021 till 30-07-2021
2ND	31-07-2021 till 29-10-2021
3RD	30-10-2021 till 28-01-2022
4TH	29-01-2022 till 29-04-2022

Table B.1: Time period of the client's fiscal year

emerge for quarter three. At the end of the 3rd fiscal quarter pick lines peak, followed by a small drop at the beginning of the 4th fiscal quarter. Lastly, it is observed that the time series appears multiplicative as the slightly upwards trend does not influence the seasonality.

In general, a time series is often tested for weekly, monthly and annual seasonality, with monthly seasonality being especially difficult because not every month has exactly the same length. This applies even more to Ceva because only business days are included. Therefore, we decided to test for weekly, quarterly and yearly seasonality instead. Hence the seasonal periods are 5, 65 and 260 respectively as there are five business days within one week, thirteen weeks of five days within a quarter and 260 business days within one year. These patterns can be tested at once, if the data is saved within R as a *multi seasonal time series* as shown in formula 2.

$$msts(Sales[, 2:6], seasonal.periods = c(5, 5*13, 5*13*4), names = names(MHE)) \qquad 2$$

Figure B.2 shows the decomposition of the trolley time series, which is a result of the 'mstl' function in R. From top to bottom the following are shown: the actual data, trend, weekly-, quarterly- and annual seasonality and finally the noise. Trolley pick lines experienced an upwards trend over the first year but the trend seems to become constant in the second year. The weekly seasonality fluctuates a lot between a highly seasonal pattern varying up to 1000 pick lines and a small seasonal pattern varying up to 500 pick

lines. Especially at the last quarter of the data, the weekly seasonality seems to be decreasing. The quarterly seasonality drops approximately 1000 lines at the end the fiscal quarter, but this drop is of very short duration. Lastly, the annual seasonality shows that the number of pick lines decreases in August and around the Christmas period. The decomposition shows that the time series contains seasonality. This highlights the importance of evaluating models, such as ETS and ARIMA, that can deal with the forecast models.
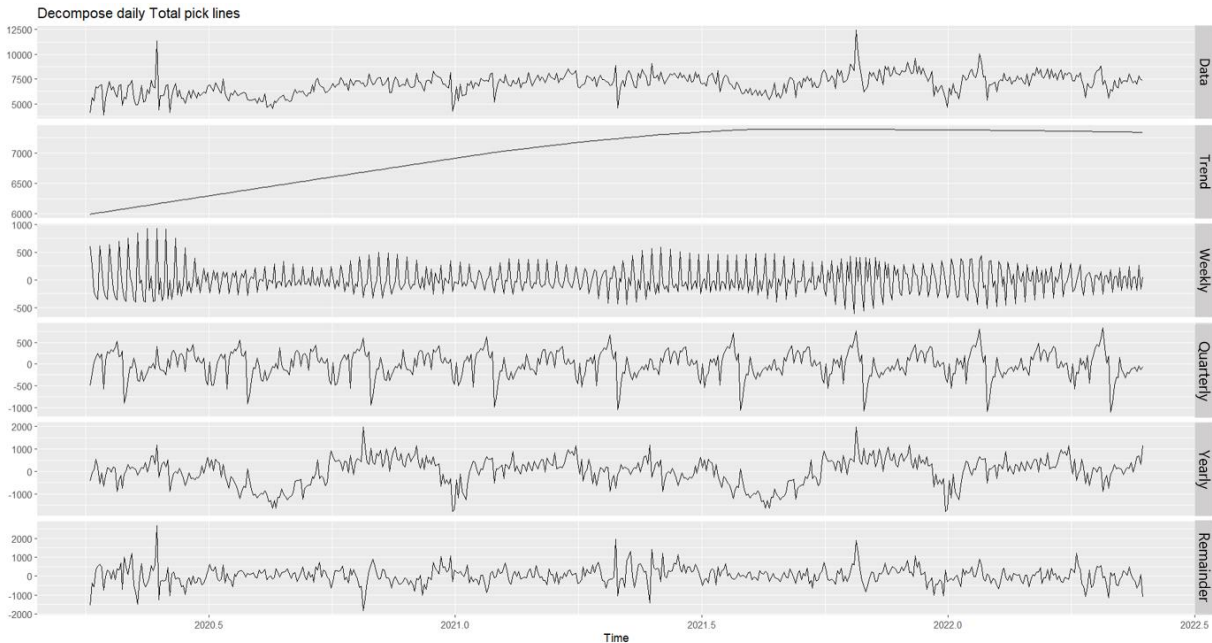
Figure B.2: Decomposition of the Trolley Pick Lines

The decomposition of the pick line time series of the other MHEs, (EPT, HOPT and Reach), shows no clear seasonality. In addition, the total amount lies around 200 pick lines per day, compared to the 7800 pick lines of trolley picks. As a result, these time series have less influence on the total pick lines per day.

Appendix C   Pseudo code ETS and ARIMA model

Algorithm 1 Pseudo Code ETS Model

1.   INPUT: Training Set Time series of each MHE
2.   OUTPUT: Forecast values and parameter values
3.
4.   *For each* MHE *do {EPT, HOPT, REACH, Trolley}*
5.   *For each* trend component *do {None, Additive, Multiplicative}*
6.   *For each* seasonal component *do {None, Additive, Multiplicative}*
7.   Determine the AIC value for the model with these parameters
8.   *If* AIC< AIC_{best} *then*
9. AIC_{best}=AIC
10.   Store the parameters of the best model so far
11.   *End if*
12.   *End for*
13.   *End for*
14.   *End for*
15.   RETURN: optimal ETS model

Algorithm 2 Pseudo Code ARIMA Model

1.   INPUT: Training set Time series of each MHE
2.   OUTPUT: Forecast values and parameter values
3.
4.   *For each* MHE *do {EPT,HOPT,REACH, Trolley}*
5.   First determine the AIC value for each of the four possible start models
6.   ARIMA(2,d,2)(1,D,1)
7.   ARIMA(0,d,0)(0,D,0)
8.   ARIMA(1,d,0)(1,D,0)
9.   ARIMA(0,d,1)(0,D,1)
10.   *Select the start model with the lowest AIC value*
11.
12.   *Now we consider variations on the current model*
13.   Vary p, q, P, Q in turns by ± 1 from the current model
14.   Vary p, q simultaneously by ± 1 from the current model
15.   Vary P, Q simultaneously by ± 1 from the current model
16.
17.   *For each* variations do
18.   Determine the AIC value
19.   *If* AIC< AIC_{best} *then*
20. AIC_{best}=AIC
21.   Store the parameters of the best model so far
22.   *End if*
23.   *End for*
24. End for
25.   RETURN: optimal ARIMA model

Appendix D   Process time trolley picking

Linear Regression of ProcessTimes dependent on the number of Pick Lines per trolley
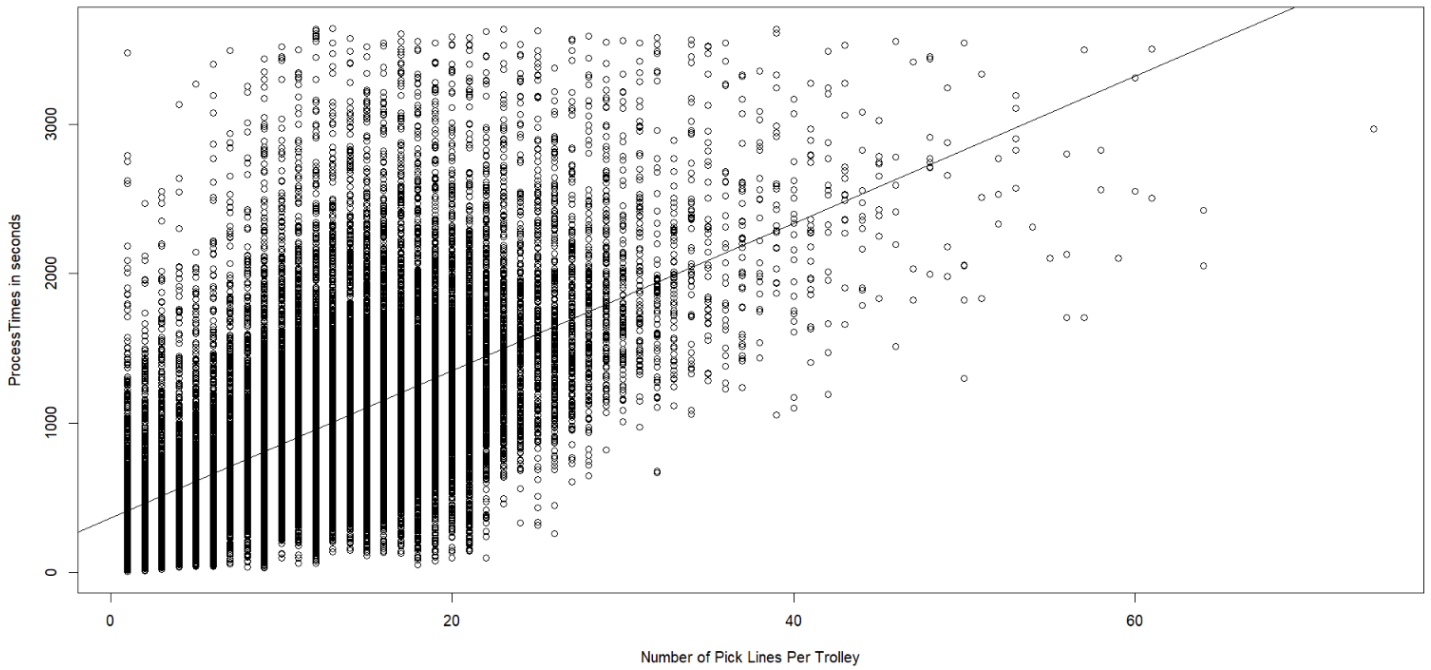

Figure D.1: Linear regression of process time dependent on the number of pick lines per trolley

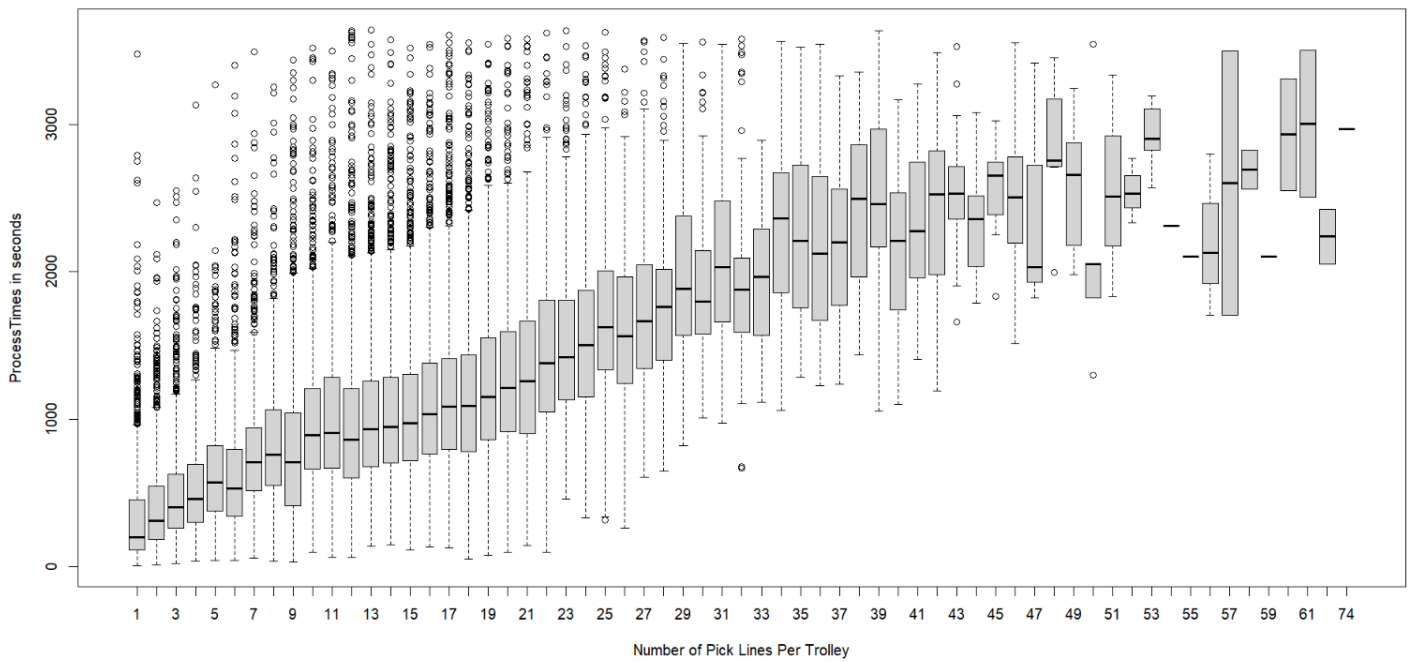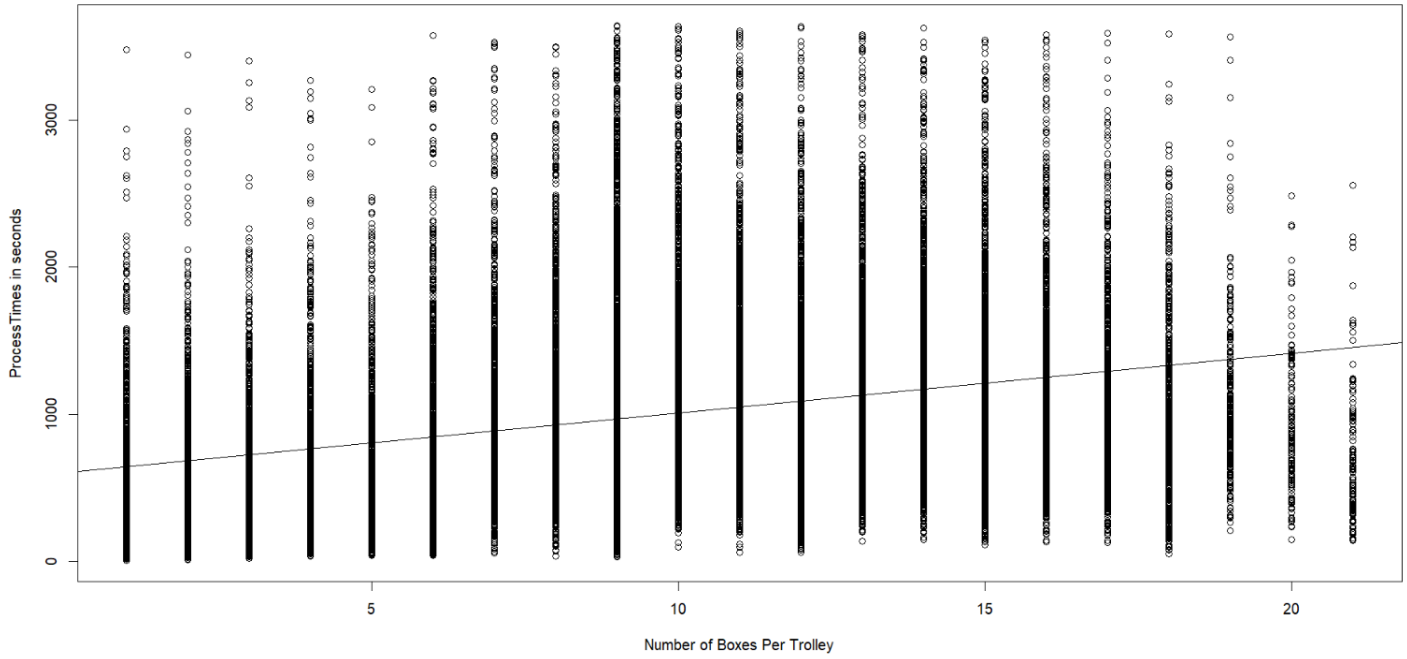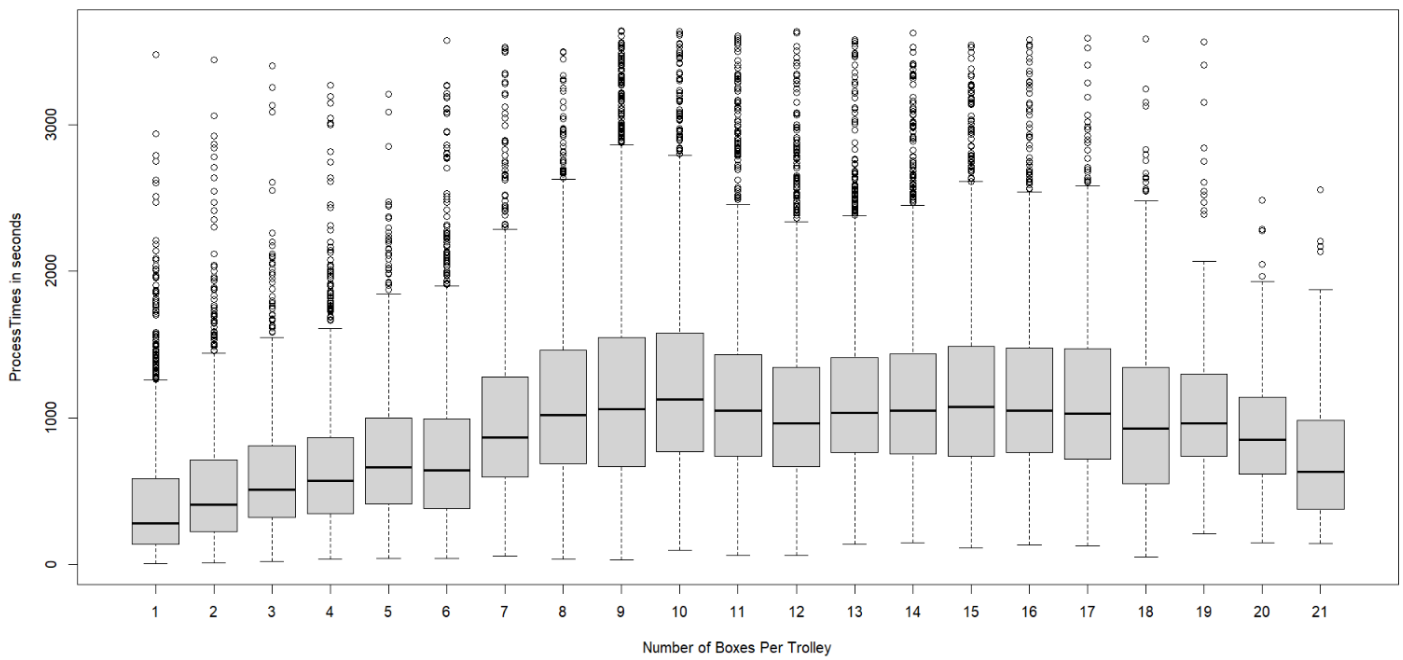**Boxplot of ProcessTimes dependent on the number of Pick Lines per trolley**

Figure D.2: Boxplot of process times dependent on the number of pick lines per trolley

Figure D.3: Linear regression of process time dependent on the number of boxes per trolley

Figure D.4: Boxplot of process times dependent on the number of boxes per trolley
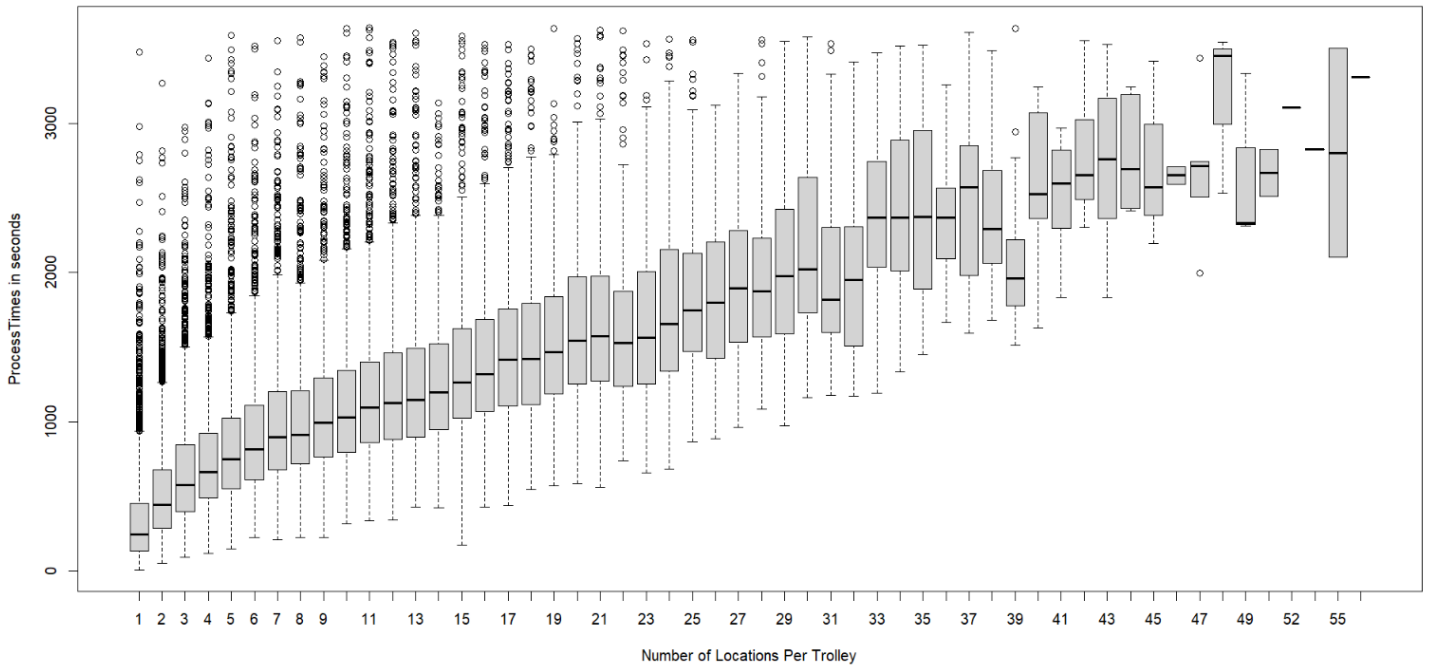
Figure D.5: Linear regression of process time dependent on the number unique locations visited per trolley
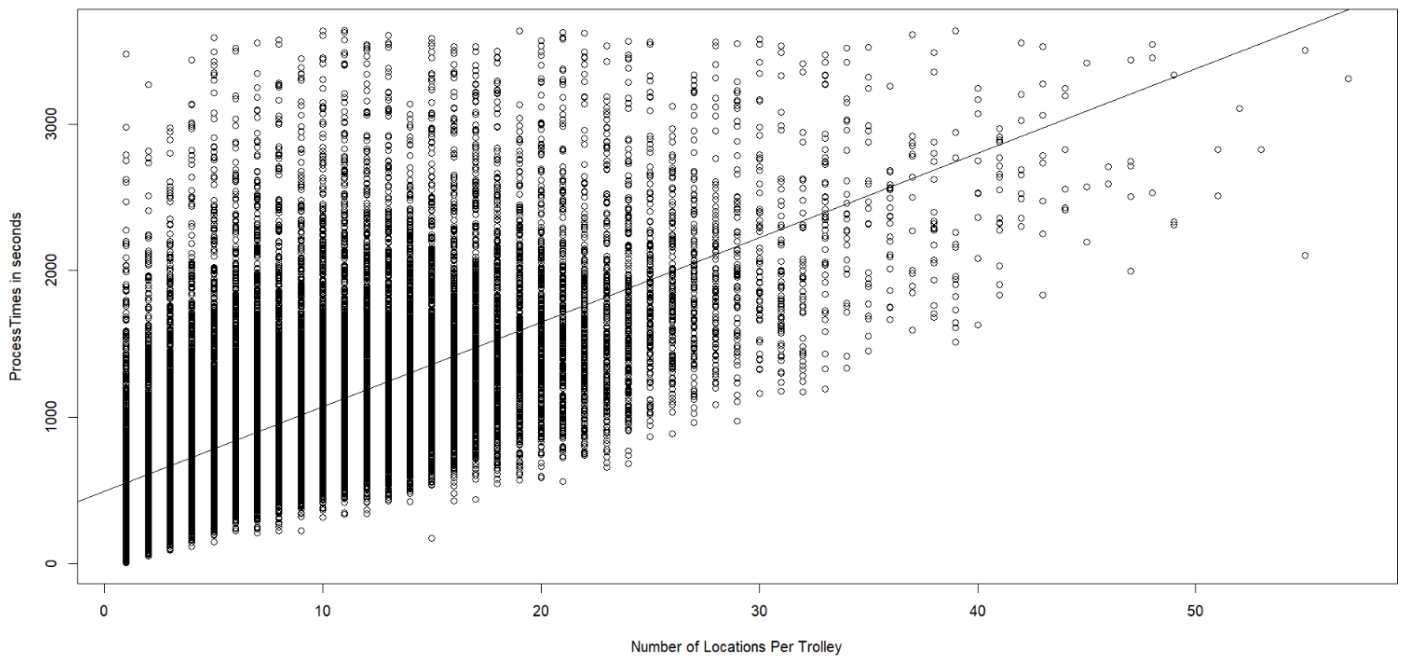

Figure D.6: Boxplot of process times dependent on the number unique locations visited per trolley

Appendix E   Grouping Pick Lines to BoxIDs and TrolleyIDs

As explained in section 2.1.3, after an order arrives it is divided into pick lines. Each pick line is a work line that contains all the required information to pick the right products, thus the item number, the quantity and location. Ceva Heerlen consists of packing machines that are capable of creating a box from a range of dimensions. The WMS system can batch multiple pick lines into a single box, by taking into account the product volume of the pick line and the box sizes which can be created by the packsize machines. The goal is to minimise the number of boxes which are sent to the end-customer as well as the empty space in the box.

A data analysis of 3.5 months is performed to determine the discrete distribution of the number pick lines per box, the results are depicted in Figure E.1. The analysis reveals that in more than 83% of the cases there is only one pick line batched to a box. For unknown orders, this discrete distribution is used in the scheduling tool to batch the pick lines to a BoxID number.
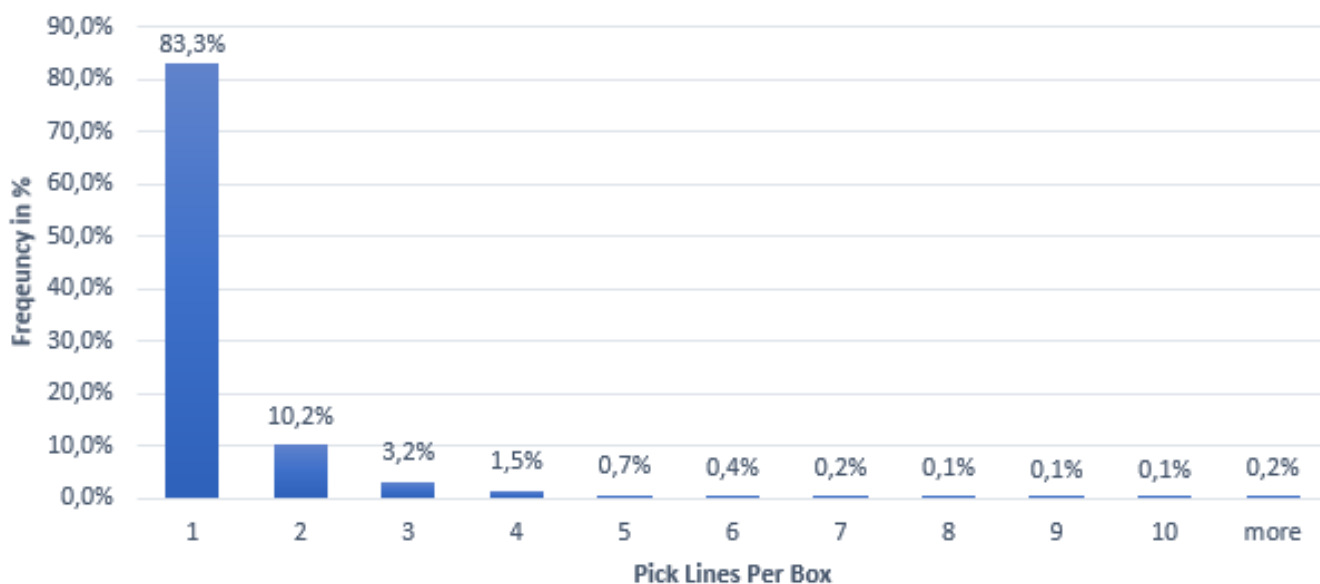

Figure E.1 Discrete distribution of pick lines per box measured over 3.5 months

Subsequently, a resource planner within Ceva called the wave planner, has to assign the boxes to a unique TrolleyID. In general, TrolleyIDs are generated based on the cut-off time and due date. Hence, each hour the waveplanner sorts the list per cut-off time and due date. Within each cut-off time and due date, the pick locations are also sorted. Based on the sizes of the boxes, the sorted boxes are added to the first TrolleyID until there is no room for an additional box. Then the next TrolleyID is created.

The wave planner usually creates around 50 TrolleyIDs simultaneously. Once the TrolleyIDs are created, the packsize machine can start manufacturing the boxes, after which the picking process can start. It is decided to create 50 trolleys at once so that on the one hand it remains well-organized at the packsize machine, but on the other hand there are sufficient pick locations to group trolleys sub-optimal. To goal is group trolleys in such a way that the walking distance of the picker remains relatively low, as walking is considered waste in a warehouse. However, if a wave planner has orders that are approaching the cut-off time then he manually overwrites the system and creates a trolley with only boxes with a specific cut-off time. In this case, a trolley is no longer sub-optimal, as the picking locations can be spread all over the warehouse.

Since the method to create trolleys depends on the manual action of a waveplanner, it is difficult to simulate the reality exactly. Therefore a data analysis of 3.5 months is performed to determine the discrete distribution of the number pick lines per box, the results are depicted in Figure E.2.
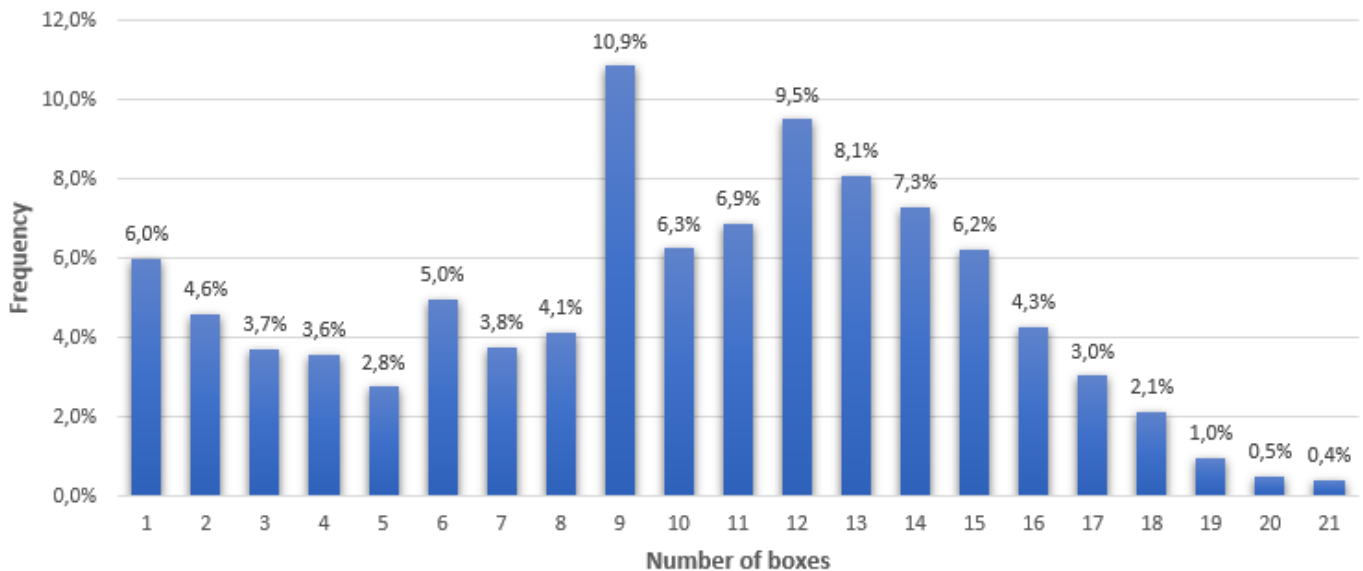

Figure E.2: Discrete distribution of the number of boxes per trolleys, measured over 3.5 months

From Figure E.2 it can be deduced that in 55% of the cases a trolley contains between 9 and 15 boxes. If a trolley contains more boxes, the dimensions of the boxes are relatively small. However, if a trolley contains fewer than 9 boxes, the trolley has been created manually. This is often due to an approaching cut-off time. TrollleyIDs with a single box are more common than expected, however if there is an issue with a box, a new TrolleyID must be often generated.

Since the trolleys with a lower number of boxes are also physically picked, the discrete distribution from Figure E.2 is used in the scheduling tool to group the BoxIDs to TrolleyIDs.

Appendix F   Ratio workflow 1,2 and 3

Trolley Pick Lines; Ratio Standard, Consolidation and Nesting
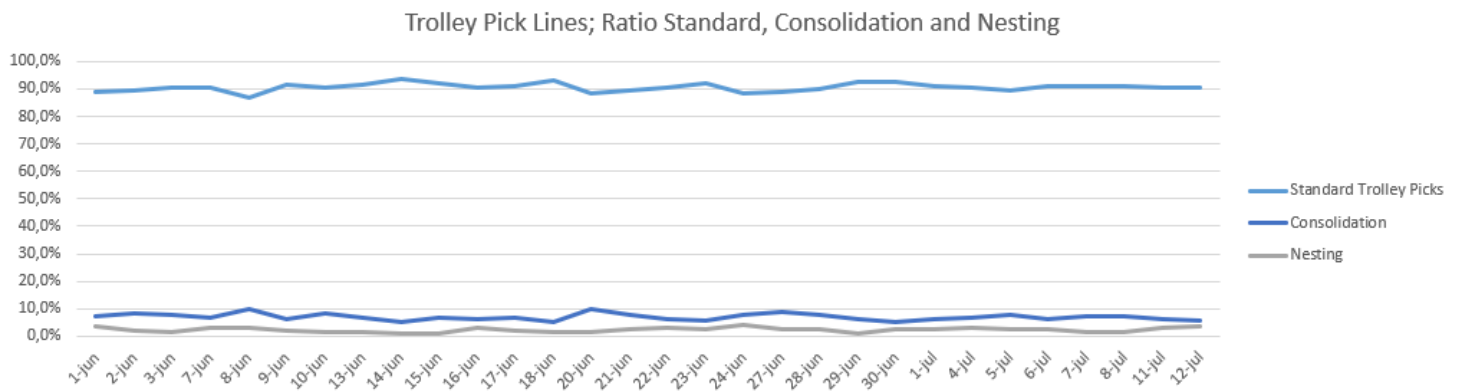
References

[1] Heerkens, H., &A. Van Winden. (2021). *Solving managerial problems systematically*: Noordhoff Uitgevers.

[2] Hans, E. W., M. v. Houdenhoven, &P. J. Hulshof. (2012). A framework for healthcare planning and control. In *Handbook of healthcare system scheduling* (pp. 303-320): Springer.

[3] Brockwell, P. J., &R. A. Davis. (2002). *Introduction to time series and forecasting*: Springer.

[4] Hyndman, R. J., &G. Athanasopoulos. (2021). *Forecasting: principles and practice* (3rd ed.): OTexts: Melbourne, Australia.

[5] Fabianová, J., P. Kačmáry, V. Molnar, &P. Michalik. (2016). Using a Software Tool in Forecasting: a Case Study of Sales Forecasting Taking into Account Data Uncertainty. *Open Engineering, 6*(1).

[6] Tompkins, J. A., J. A. White, Y. A. Bozer, &J. M. A. Tanchoco. (2010). *Facilities planning*: John Wiley & Sons.

[7] Bondarev, P. (2012). Demand Forecasting, Resource Planning and Procurement Strategy: Review and Sample Case. *Independent Analytical Research*.

[8] Luebbers, D., U. Grimmer, &M. Jarke. (2003). Systematic development of data mining-based data quality tools. In *Proceedings 2003 VLDB Conference* (pp. 548-559): Elsevier.

[9] Makridakis, S., E. Spiliotis, &V. Assimakopoulos. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PloS one, 13*(3), e0194889.

[10] Islam, B. U. (2011). Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems. *International Journal of Computer Science Issues (IJCSI), 8*(5), 504 -513

[11] Papaioannou, G. P., C. Dikaiakos, A. Dramountanis, &P. G. Papaioannou. (2016). Analysis and modeling for short-to medium-term load forecasting using a hybrid manifold learning principal component model and comparison with classical statistical models (SARIMAX, Exponential Smoothing) and artificial intelligence models (ANN, SVM): The case of Greek electricity market. *Energies, 9*(8), 635.

[12] Züfle, M., A. Bauer, V. Lesch, C. Krupitzer, N. Herbst, S. Kounev, &V. Curtef. (2019). Autonomic forecasting method selection: Examination and ways ahead. In *2019 IEEE International Conference on Autonomic Computing (ICAC)* (pp. 167-176): IEEE.

[13] Van Gils, T., K. Ramaekers, A. Caris, &M. Cools. (2017). The use of time series forecasting in zone order picking systems to predict order pickers' workload. *International Journal of Production Research, 55*(21), 6380-6393.

[14] Lalou, P., S. T. Ponis, &O. K. Efthymiou. (2020). Demand forecasting of retail sales using data analytics and statistical programming. *Management & Marketing. Challenges for the Knowledge Society, 15*(2), 186-202.

[15] Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting, 20*(1), 5-10.

[16] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management science, 6*(3), 324-342.

[17] Gardner Jr, E. S., &E. McKenzie. (1985). Forecasting trends in time series. *Management science, 31*(10), 1237-1246.

[18] Hyndman, R. J., A. B. Koehler, R. D. Snyder, &S. Grose. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of forecasting, 18*(3), 439-454.

[19] Assimakopoulos, V., &K. Nikolopoulos. (2000). The theta model: a decomposition approach to forecasting. *International journal of forecasting, 16*(4), 521-530.

[20]   Hyndman, R. J., &B. Billah. (2003). Unmasking the Theta method. *International Journal of Forecasting, 19*(2), 287-290.

[21]   Dudek, G. (2019). Short-term load forecasting using Theta method. In *E3S Web of Conferences* (Vol. 84, pp. 01004): EDP Sciences.

[22]   Chen, P., A. Niu, D. Liu, W. Jiang, &B. Ma. (2018). Time series forecasting of temperatures using SARIMA: An example from Nanjing. In *IOP Conference Series: Materials Science and Engineering* (Vol. 394, pp. 052024): IOP Publishing.

[23]   Mohri, M., A. Rostamizadeh, &A. Talwalkar. (2018). *Foundations of machine learning*: MIT press.

[24]   Adhikari, R., &R. K. Agrawal. (2013). An introductory study on time series modeling and forecasting. *arXiv preprint arXiv:1302.6613*.

[25]   Fattah, J., L. Ezzine, Z. Aman, H. El Moussami, &A. Lachhab. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management, 10*, 1847979018808673.

[26]   Berrar, D. (2019). Cross-Validation. In.

[27]   Groover, M. P. (2007). *Work systems and the methods, measurement, and management of work*: Pearson Prentice Hall Upper Saddle River, NJ.

[28]   Law, A. M., W. D. Kelton, &W. D. Kelton. (2007). *Simulation modeling and analysis* (Vol. 3): Mcgraw-hill New York.

[29]   Van den Bergh, J., J. Beliën, P. De Bruecker, E. Demeulemeester, &L. De Boeck. (2013). Personnel scheduling: A literature review. *European journal of operational research, 226*(3), 367-385.

[30]   Muter, İ., &T. Öncan. (2022). Order batching and picker scheduling in warehouse order picking. *IISE Transactions, 54*(5), 435-447.

[31]   Castillo-Salazar, J. A., D. Landa-Silva, &R. Qu. (2016). Workforce scheduling and routing problems: literature survey and computational study. *Annals of Operations Research, 239*(1), 39-67.

[32]   Pinedo, M., &K. Hadavi. (1992). Scheduling: theory, algorithms and systems development. In *Operations research proceedings 1991* (pp. 35-42): Springer.

[33]   Baykasoglu, A. (2002). Linguistic-based meta-heuristic optimization model for flexible job shop scheduling. *International Journal of Production Research, 40*(17), 4523-4543.

[34]   Müller, D. (2021). Scheduling flexible job shops under workforce constraints.

[35]   Fattahi, P., M. Saidi Mehrabad, &F. Jolai. (2007). Mathematical modeling and heuristic approaches to flexible job shop scheduling problems. *Journal of intelligent manufacturing, 18*(3), 331-342.

[36]   Ham, A. M., &E. Cakici. (2016). Flexible job shop scheduling problem with parallel batch processing machines: MIP and CP approaches. *Computers & Industrial Engineering, 102*, 160-165.

[37]   Andrade-Pineda, J. L., D. Canca, P. L. Gonzalez-R, &M. Calle. (2020). Scheduling a dual-resource flexible job shop with makespan and due date-related criteria. *Annals of Operations Research, 291*(1), 5-35.

[38]   Topaloglu, S., &G. Kilincli. (2009). A modified shifting bottleneck heuristic for the reentrant job shop scheduling problem with makespan minimization. *The International Journal of Advanced Manufacturing Technology, 44*(7), 781-794.

[39]   Chen, B., &T. I. Matis. (2013). A flexible dispatching rule for minimizing tardiness in job shop scheduling. *International Journal of Production Economics, 141*(1), 360-365.

[40]   Yamada, T., &R. Nakano. (1997). Genetic algorithms for job-shop scheduling problems. *Proceedings of modern heuristic for decision support*, 67-81.

[41]   Heger, J., S. Grundstein, &M. Freitag. (2017). Online-scheduling using past and real-time data. An assessment by discrete event simulation using exponential smoothing. *CIRP Journal of Manufacturing Science and Technology, 19*, 158-163.

[42]   Huber, J., &H. Stuckenschmidt. (2020). Daily retail demand forecasting using machine learning with emphasis on calendric special days. *International Journal of Forecasting, 36*(4), 1420-1438.

[43]   Hyndman, R. J. (2010). Forecasting with long seasonal periods. Retrieved from https://robjhyndman.com/hyndsight/longseasonality/

[44]   Hyndman, R. J., &Y. Khandakar. (2008). Automatic time series forecasting: the forecast package for R. *Journal of statistical software, 27*, 1-22.

UNIVERSITY OF TWENTE.